

PROJECTIVELY FLAT FINSLER 2-SPHERES OF CONSTANT CURVATURE

ROBERT L. BRYANT

Duke University

November 25, 1996

ABSTRACT. After recalling the structure equations of Finsler structures on surfaces, I define a notion of ‘generalized Finsler structure’ as a way of micro-localizing the problem of describing Finsler structures subject to curvature conditions. I then recall the basic notions of path geometry on a surface and define a notion of ‘generalized path geometry’ analogous to that of ‘generalized Finsler structure’.

I use these ideas to study the geometry of Finsler structures on the 2-sphere that have constant Finsler-Gauss curvature K and whose geodesic path geometry is projectively flat, i.e., locally equivalent to that of straight lines in the plane.

I show that modulo diffeomorphism there is a 2-parameter family of projectively flat Finsler structures on the sphere whose Finsler-Gauss curvature K is identically 1.

0. INTRODUCTION

Hilbert’s Fourth Problem was entitled “Problem of the straight line as the shortest distance between two points”. It concerned, in its most general formulation, the problem of characterizing the not-necessarily-symmetric distance functions d that could be defined on (convex) subsets $U \subset \mathbb{R}^2$ so that the lines were geodesics, i.e., so that $d(x, z) \leq d(x, y) + d(y, z)$ with equality if and only if x , y , and z are collinear, with y lying on the segment joining x to z .

Hilbert’s reason for considering non-symmetric distances was that interesting non-symmetric examples had already been discovered by Minkowski. He was also aware that notions of length of curves defined in many calculus of variations problems leads naturally to non-symmetric distance functions. For some interesting examples of physical relevance, see Carathéodory’s book [Cara].

The Fourth Problem can be regarded as a fundamental example of the inverse problem in the calculus of variations. That is, given that the straight lines are the extremals (i.e., geodesics) of a first order Lagrangian for oriented curves in the

1991 *Mathematics Subject Classification.* 53C60, 53A20, 58G30.

Key words and phrases. Finsler geometry, projective geometry.

Research for this article was begun while I was visiting IMPA in Rio de Janeiro in August 1996, while the actual writing took place while I was visiting the IAS in Princeton in the Fall of 1996. Their hospitality as well as NSF support through the grant DMS-9505125 is hereby gratefully acknowledged.

plane, what can one say about the Lagrangian? Can the set of such Lagrangians be characterized in some useful way? Can their stability properties be understood?

Beltrami had shown in 1866 that any Riemannian metric on an open subset of the plane whose geodesics are the straight lines must have constant Gauss curvature. In fact, such a metric must be locally equivalent via a projective transformation to one of the standard metrics of constant curvature defined on the plane or an open subset thereof, such as the Klein model of the hyperbolic plane as the unit disk, with the geodesics being the chords of the boundary circle. Thus, the problem was already solved for Lagrangians that represented Riemannian arc length.

More general Lagrangians for curves on a domain M in the plane (or, more generally, on any surface) can be described in terms of Finsler norms, where a *Finsler norm* or *metric* is a non-negative function $L : TM \rightarrow \mathbb{R}$ that is positive and smooth away from the zero section and has the *homogeneity property* that $L(x, \lambda v) = \lambda L(x, v)$ for all $\lambda \geq 0$ and all $v \in T_x M$ as well as the *convexity property* that the unit sphere (or, as it is classically known, the *indicatrix*)

$$\Sigma_x = \{v \in T_x M \mid L(x, v) = 1\} \subset T_x M$$

at $x \in M$ be a smooth, closed, strictly convex curve in $T_x M$ for all $x \in M$. If, in addition, $L(x, -v) = L(x, v)$, or, equivalently, $\Sigma_x = -\Sigma_x$, then L is said to be *symmetric*.

The homogeneity property implies that for any immersed, oriented curve $\gamma : [a, b] \rightarrow M$ the integral

$$\mathcal{L}(\gamma) = \int_a^b L(\gamma(t), \gamma'(t)) dt$$

is unchanged by oriented reparametrization and so defines a notion of L -length for oriented curves in M . The oriented curves that are extremal for the functional \mathcal{L} are known as *L -geodesics*. When L is symmetric, the L -length of a curve is independent of its orientation.

The convexity property implies that the Euler-Lagrange equations for the geodesics of L are everywhere non-degenerate, so that, in each oriented direction through each point of M there will pass a unique L -geodesic and these L -geodesics will depend smoothly on parameters, just as in the Riemannian case.

The most familiar example of a Finsler norm is $L = \sqrt{g}$ where g is a Riemannian metric on M . In that case, the first geometric invariant that can be attached to L after the geodesics themselves is the Gauss curvature K , which is a function on M . Besides being a diffeomorphism invariant of the metric g , it enters fundamentally into the study of the second variation of the geodesics: For any unit speed geodesic $\gamma : [a, b] \rightarrow M$, the Jacobi equation for variation of geodesics near γ is

$$y''(s) + K(\gamma(s)) y(s) = 0.$$

For example, γ is locally minimizing when the operator $Jy = y'' + K \circ \gamma y$ has zero nullity and index.

In the more general Finsler case, the second variation formula still makes sense, but now K has to be defined, not on M , but on the unit sphere bundle $\Sigma \subset TM$. Then, for a unit speed geodesic γ as above, the Jacobi equation turns out to be

$$y''(s) + K(\gamma'(s)) y(s) = 0,$$

and it bears the same relation to minimizing properties of L -geodesics as the classical Jacobi equation in the Riemannian case. This form of the second variation was first discovered in 1907 by Underhill [Un]. The function K (to be described more explicitly in §1 below) was given the name ‘inneres Krümmungsmaß’ by Finsler in his 1918 dissertation on the geometry of Finsler spaces, but in the present article, it will be called the *Finsler-Gauss curvature*.

It turned out that there were many solutions to Hilbert’s Fourth Problem, and the interested reader can profitably consult the survey article [Bu] or the book [Po], where a complete solution in the case of symmetric distances is presented. Motivated by the central place that metrics of constant curvature play in Riemannian geometry and the direct relation that it bears with the stability question, it is natural to pay special attention to the Finsler norms for which K is constant.

In 1929, Funk [Fu1] showed that the Finsler examples constructed by Hilbert on arbitrary convex domains in the plane satisfied $K = -1$ and went on to classify all such solutions to Hilbert’s Fourth Problem on convex domains in the plane. He also classified the solutions with $K = 0$. In each case, the local solutions depended on arbitrary functions of one variable. In 1935 [Fu2], he returned to the subject and found a local formula for the solutions to Hilbert’s Fourth Problem that satisfied $K = +1$. He showed that the local solutions could be described in terms of an arbitrary holomorphic function of one variable subject to some inequalities. In 1963 [Fu3], he showed that the only global solutions on the entire projective plane that satisfied certain other local conditions, including symmetry, were the Riemannian solutions.

There remains the question of whether there are any non-Riemannian solutions to Hilbert’s Fourth Problem defined on the entire projective plane or the sphere and satisfying $K = 1$. In his 1976 survey article, Busemann [Bu, p. 139] wrote “... Funk [determined] all two-dimensional Desarguesian metrics with constant positive curvature. Even [these] depend on an essentially arbitrary function, where ‘essential’ means that the function is restricted by an equality only.” (In this context, ‘Desarguesian metric’ can be taken to mean ‘Finsler metric with linear geodesics’.) From the context of Busemann’s remark, it seems clear that he thought that the Finsler metrics on the 2-sphere having constant curvature and solving Hilbert’s Fourth Problem would depend on an ‘arbitrary’ function.

In a 1986 paper providing a new solution to Hilbert’s Fourth Problem in the symmetric case, Szabó states [Sz, pp. 297–299] that he will discuss the non-symmetric case with constant positive curvature in the sequel, but, to my knowledge, this sequel has never appeared.

In this article, I show that there are indeed non-Riemannian global solutions to Hilbert’s Fourth Problem having constant curvature $K = 1$. Despite Busemann’s claim, I further show that, up to diffeomorphism, there is exactly a 2-parameter family of inequivalent solutions and that the only symmetric solution is the Riemannian one. An explicit formula for these Finsler metrics is given in Theorem 10. The unit sphere or indicatrix in these examples turns out to be a smooth algebraic plane curve of degree 4 (or, at some special points or in the Riemannian case, 2) and the only symmetric Finsler structure among them is the Riemannian one.

This article is organized as follows: In §1, I recall the basics of the geometry of Finsler structures and introduce a generalization that will be needed in the rest of the article. This generalization allows for multi-valued Finsler norms (see Example 1) and other apparent pathologies, such as Finsler structures on orbifolds,

but it has the virtue that the differential equations for prescribed curvature in Finsler geometry are local on the generalized objects, while they are not on the classical objects. This generalization is actually of independent interest and may have applications to control theory.

In §2, I go over the basics of path geometry, i.e., the geometry of a surface endowed with a 2-parameter family of paths, such as the geodesic path geometry of a Finsler metric. I introduce a notion of generalized path geometry that corresponds to the earlier notion of generalized Finsler structure and use this to develop a notion of and a test for projective flatness for a generalized Finsler structure. I recall Cartan's projective connection for a path geometry and note that his construction extends without change to the case of a generalized path geometry. His computation of projective curvatures is used to recover Berwald's characterization of the Finsler structures whose geodesics can be mapped to straight lines in the plane by local diffeomorphism (the projectively flat case) and to extend this concept to the generalized case.

In §3, I classify the projectively flat Finsler structures with $K = 1$ that can be defined on compact simply-connected 3-manifolds. I show that up to diffeomorphism, there is a 2-parameter family of these and that they each arise as the double cover of the unit sphere bundle of a classical Finsler structure on the 2-sphere. The main tool in the finiteness theorem is a vanishing theorem for a certain holomorphic cubic differential on a Riemann surface of genus zero. The existence theorem uses the vanishing theorem to provide enough extra local equations to identify the space of solutions with the leaves of an integrable distribution of dimension 3 defined on a certain manifold X of dimension 13.

The constructions of §3 are invariant under the projective group acting on \mathbb{RP}^2 and they suggest that a projectively invariant interpretation of Funk's local characterization of these Finsler structures in the classical case might lead to a global classification theorem. In §4, I do just this. I recast Funk's results of [Fu1-3] as statements about sections of a certain bundle over the double cover of \mathbb{RP}^2 and the geometry of certain holomorphic curves in \mathbb{CP}^2 . Global considerations from algebraic geometry make it possible to then classify the holomorphic curves that can arise in any global solution as the conics without real points. Then, using a classification of the moduli of conics without real points under the action of the real group $SL(3, \mathbb{R})$, I derive the explicit formulae of Theorem 10. I then close the paper with a few remarks about the geometry of the examples.

Throughout the article, smoothness is assumed. Weaker differentiability assumptions would have sufficed, but would have been a distraction from the main arguments. It appears likely that the classification results would hold without change as long as the Finsler structures were differentiable enough to define the curvature, but I have not pursued this question.

1. FINSLER SURFACES AND GENERALIZED FINSLER STRUCTURES

This section contains a brief review of basic Finsler surface theory and Cartan's construction of the canonical coframing associated to a Finsler structure on a surface. Along with this review, I will also introduce and discuss a notion of generalized Finsler structure that will be used in the rest of the article.

The material on classical Finsler surfaces has been treated in many places, cf. [BaChSh], [Ca2], [Ch1], [Ru], [Ma2], and [GaWi], to name just a few from

different points of view and different eras. It is included here to establish notation and nomenclature.

Let M be a connected, smooth, oriented surface (i.e., a 2-manifold). A *Finsler structure* on M is a smooth hypersurface $\Sigma^3 \subset TM$ for which the basepoint projection $\pi : \Sigma \rightarrow M$ is a surjective submersion and having the property that for each $x \in M$, the π -fiber $\Sigma_x = \pi^{-1}(x) = \Sigma \cap T_x M$ is a closed, strictly convex curve enclosing the origin $0_x \in T_x M$. If, in addition, the curve Σ_x is symmetric about 0_x for each $x \in M$, then Σ will be said to be *symmetric*.

A differentiable curve $\gamma : [a, b] \rightarrow M$ will be said to be a Σ -curve if, for every s in the interval $[a, b]$, the velocity vector $\gamma'(s)$ lies in Σ . The map $\gamma' : [a, b] \rightarrow \Sigma$ is known as the *tangential lift* of γ . For every immersed curve $\gamma : [a, b] \rightarrow M$, there is a unique orientation preserving diffeomorphism $h : [0, L] \rightarrow [a, b]$ so that $\gamma \circ h$ is a Σ -curve. The number $L > 0$ is the Σ -length of γ . Note that if Σ is not symmetric then reversing the orientation of an immersed curve may change its length.

1.1. The canonical coframing. In [Ca2], É. Cartan constructed a canonical coframing on any Finsler structure $\Sigma \subset TM$ on an oriented surface M . In that paper, he essentially proves the following result.

Theorem 1 (Cartan). *Let $\Sigma \subset TM$ be a Finsler structure on an oriented surface M and let $\pi : \Sigma \rightarrow M$ denote the basepoint projection. Then there exists a unique coframing $\omega = (\omega_1, \omega_2, \omega_3)$ of Σ with the properties:*

- (1) $\omega_1 \wedge \omega_2$ is a positive multiple of any π -pullback of a positive 2-form on M ;
- (2) The tangential lift of any Σ -curve γ satisfies $(\gamma')^* \omega_2 = 0$ and $(\gamma')^* \omega_1 = ds$;
- (3) $\omega_2 \wedge d\omega_1 = 0$;
- (4) $\omega_1 \wedge d\omega_1 = \omega_2 \wedge d\omega_2$;
- (5) $d\omega_1 = -\omega_2 \wedge \omega_3$ and $\omega_3 \wedge d\omega_2 = 0$.

Moreover, there exist unique functions I , J , and K on Σ so that

$$(1) \quad \begin{aligned} d\omega_1 &= -\omega_2 \wedge \omega_3, \\ d\omega_2 &= -\omega_3 \wedge (\omega_1 - I \omega_2), \\ d\omega_3 &= -(K \omega_1 - J \omega_3) \wedge \omega_2. \end{aligned}$$

Let $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$ be the vector field framing of Σ that is dual to the coframing $\omega = (\omega_1, \omega_2, \omega_3)$. In terms of \mathbf{X} , equations (1) can be expressed in the form

$$(2) \quad \begin{aligned} [\mathbf{X}_2, \mathbf{X}_3] &= \mathbf{X}_1 + I \mathbf{X}_2 + J \mathbf{X}_3, \\ [\mathbf{X}_3, \mathbf{X}_1] &= \mathbf{X}_2, \\ [\mathbf{X}_1, \mathbf{X}_2] &= K \mathbf{X}_3. \end{aligned}$$

Note that the fibers of the basepoint projection π are simply the integral curves of the vector field \mathbf{X}_3 .

Remarks. The notation is Cartan's. However, the reader should be aware that other authors have modified it somewhat. For example, H. Rund denotes the invariant I by the letter J . In a recent paper [Br], I used the letters S and C to denote the quantities $-I$ and $-J$, a decision that I now regret.

Strictly speaking, there is a slight difference between Theorem 1 and the result of Cartan. Since Cartan does not fix an orientation on the surface M , the first property in the list above has no meaning for him. As a result, Cartan's coframing is only well-defined up to an ambiguity of the form $\omega = (\omega_1, \pm\omega_2, \pm\omega_3)$. The introduction of an orientation on M fixes this ambiguity.

In standard treatments of the calculus of variations, the 1-form ω_1 is known as *Hilbert's invariant integral*.

A Σ -curve γ is a Σ -geodesic, i.e., a critical point of the natural Σ -length functional on curves in M , if and only if its tangential lift satisfies $(\gamma')^*\omega_3 = 0$. Thus, the Σ -geodesics are the projections to M of the integral curves of \mathbf{X}_1 . For this reason, the flow of the vector field \mathbf{X}_1 is known as the *geodesic flow* of Σ and Σ is said to be *geodesically complete* if \mathbf{X}_1 is complete, i.e., its flow from any initial point is defined for all time (positive and negative).

The function I vanishes if and only if Σ is the unit circle bundle of a Riemannian metric g on M . In this case, the function K is the π -pullback to Σ of the Gauss curvature function of g .

1.2. Generalized Finsler structures. Taking Cartan's construction as a starting point leads to a natural widening of the notion of a Finsler structure.

Definition 1. A *generalized Finsler structure* on a 3-manifold Σ is a coframing $\omega = (\omega_1, \omega_2, \omega_3)$ that satisfies the equations (1) for some (necessarily unique) functions I , J , and K on Σ .

As in the classical case, for a generalized Finsler structure $\omega = (\omega_1, \omega_2, \omega_3)$, the dual framing of vector fields will be denoted $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$.

A generalized Finsler structure will be said to be *amenable* if the leaf space of the foliation defined by the integral curves of \mathbf{X}_3 can be given the structure of a smooth surface M in such a way that the natural projection $\pi : \Sigma \rightarrow M$ is a smooth submersion.

Every generalized Finsler structure (Σ, ω) is locally amenable in the sense that every point of Σ has a neighborhood to which the generalized Finsler structure restricts to be amenable. In fact, the next proposition shows that the difference between the concepts 'Finsler structure' and 'generalized Finsler structure' is global in nature; every generalized Finsler structure is locally diffeomorphic to a Finsler structure. The proof is straightforward, so I omit it.

Proposition 1. *Let Σ be a 3-manifold endowed with an amenable generalized Finsler structure ω . Denote the projection onto the space of integral curves of \mathbf{X}_3 by $\pi : \Sigma \rightarrow M$ and define a smooth map $\nu : \Sigma \rightarrow TM$ by the rule $\nu(\mathbf{u}) = \pi'(\mathbf{u})(\mathbf{X}_1(\mathbf{u}))$ for all $\mathbf{u} \in \Sigma$. Then ν immerses each π -fiber $\Sigma_x = \pi^{-1}(x)$ as a curve in $T_x M$ that is strictly convex towards 0_x . Moreover, there is an orientation of M so that the ν -pullback of the canonical coframing induced on the ν -image of Σ coincides with the given generalized Finsler structure. \square*

The reader may well wonder why anyone would bother with generalized Finsler structures since they are locally the same as Finsler structures. The reason is that in the study of Finsler structures defined by geometric conditions, such as conditions on the invariants I , J , and K , one is frequently led to solve differential equations in the larger class of generalized Finsler structures since it is this class that is locally defined. Then, as a separate step, one can determine the necessary conditions

on a generalized Finsler structure that it actually be a Finsler structure. Thus, generalized Finsler structures provide a natural intermediate stage where problems can be localized and solved without the complication of global issues.

The next proposition gives a simple necessary and sufficient test for a generalized Finsler structure to be a Finsler structure. The proof is straightforward.

Proposition 2. *A generalized Finsler structure ω on a 3-manifold Σ is a Finsler structure if and only if the integral curves of \mathbf{X}_3 are all closed, it is amenable, and the canonical immersion $\nu : \Sigma \rightarrow TM$ is one-to-one. \square*

Note that just having the integral curves of \mathbf{X}_3 be closed does not make a generalized Finsler structure amenable since one could have a discrete subset of the leaves around which the foliation is not locally a product, the so-called ‘ramified’ orbits. In this case, the leaf space will have the structure of a 2-dimensional orbifold near the ramified points.

When a generalized Finsler structure is amenable in addition to having all the integral curves of \mathbf{X}_3 closed, the ν -image of each π -fiber Σ_x will in general be a closed, strictly convex curve in $T_x M$ which winds around the origin μ times for some positive integer μ which may well be greater than 1. The number μ , known as the *multiplicity* of the generalized Finsler structure, is equal to 1 if and only if each Σ_x is embedded via ν .

Example 1. Let S^2 be given its standard Riemannian metric g_0 and let $\Sigma_0 \subset TS^2$ be its unit tangent bundle, with basepoint projection $\pi_0 : TS^2 \rightarrow S^2$. Of course, Σ_0 defines a Finsler structure on S^2 in the classical sense.

Now, it is well known that Σ_0 is diffeomorphic to $SO(3)$ and hence has \mathbb{Z}_2 as its fundamental group. Let $\tau : \Sigma \rightarrow \Sigma_0$ be its universal cover, so that Σ is diffeomorphic to S^3 . Now let $\tilde{\tau} : \Sigma \rightarrow TS^2$ be an arbitrary small perturbation of τ and define $\pi = \pi_0 \circ \tilde{\tau}$. Provided this perturbation is sufficiently small in the C^2 -topology, the $\tilde{\tau}$ -image of each fiber $\Sigma_x = \pi^{-1}(x) \subset \Sigma$ will be a closed, immersed curve in $T_x S^2$ which is strictly convex towards 0_x . For the generic such perturbation, these image curves will have winding number 2 about 0_x but will not double cover an embedded curve of winding number 1 about 0_x .

Now, Cartan’s construction of his canonical coframing on a hypersurface in TM is local and depends only on the assumptions that the hypersurface submerses onto M with fibers that are locally strictly convex towards the origin in each tangent space. It follows that the immersion $\tilde{\tau}$ induces a generalized Finsler structure $\tilde{\omega}$ on Σ . This generalized Finsler structure is amenable, with $\pi : \Sigma \rightarrow S^2$ being the leaf projection and $\tilde{\tau}$ being the canonical immersion ν . In this case, the multiplicity μ is equal to 2.

1.3. The Bianchi identities. The structure equations of a generalized Finsler structure ω on a 3-manifold Σ are

$$\begin{aligned} d\omega_1 &= -\omega_2 \wedge \omega_3, \\ d\omega_2 &= -\omega_3 \wedge (\omega_1 - I\omega_2), \\ d\omega_3 &= -(K\omega_1 - J\omega_3) \wedge \omega_2, \end{aligned}$$

where I , J , and K are smooth functions on Σ . Still following Cartan, I will now derive the Bianchi identities, of the structure, i.e., the relations among the derivatives of the invariants I , J , and K .

The exterior derivative of the first of these three equations is an identity when the other two are taken into account. The exterior derivative of the second equation simplifies to

$$0 = d(d\omega_2) = -(dI - J\omega_1) \wedge \omega_2 \wedge \omega_3,$$

while the exterior derivative of the third equation simplifies to

$$0 = d(d\omega_3) = (-dK \wedge \omega_1 + (dJ + KI\omega_1) \wedge \omega_3) \wedge \omega_2.$$

These two equations imply that there exist functions $I_2, I_3, J_2, J_3, K_1, K_2$, and K_3 for which

$$(3) \quad \begin{aligned} dI &= J\omega_1 + I_2\omega_2 + I_3\omega_3, \\ dJ &= -(K_3 + KI)\omega_1 + J_2\omega_2 + J_3\omega_3, \\ dK &= K_1\omega_1 + K_2\omega_2 + K_3\omega_3. \end{aligned}$$

The formulae (3) constitute the Bianchi identities of the structure. Alternatively, they can be expressed as follows: For any differentiable function F on Σ , define the functions F_1, F_2 , and F_3 by the equation $dF = F_1\omega_1 + F_2\omega_2 + F_3\omega_3$. Then the Bianchi identities can also be expressed in the form

$$I_1 - J = 0 \quad \text{and} \quad J_1 + K_3 + KI = 0.$$

1.4. Geodesics and the second variation. Let $\omega = (\omega_1, \omega_2, \omega_3)$ be a generalized Finsler structure on a 3-manifold Σ . The *geodesics* of the structure are the integral curves of \mathbf{X}_1 . They define a foliation of Σ called the *geodesic foliation*. The (local) flow of \mathbf{X}_1 is the *geodesic flow* and ω is said to be *geodesically complete* if \mathbf{X}_1 is complete, i.e., its flow exists for all time. The structure ω is *geodesically amenable* if the leaf space Λ of the geodesic foliation can be given the structure of a smooth surface in such a way that the natural projection $\ell : \Sigma \rightarrow \Lambda$ is a smooth submersion. Of course, every generalized Finsler structure is locally geodesically amenable.

When $\Sigma \subset TM$ is a Finsler structure on a surface M , a Σ -geodesic is a Σ -curve $\gamma : D \rightarrow M$ (where $D \subset \mathbb{R}$ is some interval, which may be bounded or unbounded and closed or open) for which the lifted curve $\gamma' : D \rightarrow \Sigma$ is a geodesic of the generalized Finsler structure on Σ . If $\gamma : D \rightarrow M$ is a Σ -geodesic and the interval D contains 0 then $\gamma'(s) = \exp_{s\mathbf{X}_1}(\mathbf{u})$ where $\mathbf{u} = \gamma'(0)$. Thus, $\gamma(s) = \pi(\exp_{s\mathbf{X}_1}(\mathbf{u}))$.

In the Riemannian geometry of surfaces, the Gaussian curvature plays an important role in the formula for the second variation of arc length. In the more general case of a Finsler structure, the function K , known as the Finsler-Gauss curvature, plays the same role. More precisely, a compact Σ -geodesic $\gamma : [a, b] \rightarrow M$ will be a local minimum of the Σ -length functional if the quadratic form

$$Q_\gamma(f) = \int_a^b ((f')^2 - (K \circ \gamma') f^2) dt$$

has zero index and nullity on the space of smooth functions f on $[a, b]$ which vanish at the endpoints. This follows by an elementary calculation which I omit.

In particular, if K is non-positive then every geodesic segment is locally minimizing. On the other hand, if $K \geq a^2$ for some positive constant a , then, just as in

the Riemannian case, no geodesic segment of length greater than π/a can be locally minimizing.

Remark. When a Finsler structure $\Sigma \subset TM$ is geodesically amenable, Crofton's Formula holds in the following sense: If $\ell : \Sigma \rightarrow \Lambda$ is the submersion onto the leaf space of the geodesic flow, then there exists a unique 2-form μ on Λ so that $\ell^* \mu = \omega_2 \wedge \omega_3$. For any immersed curve $\gamma : [0, 1] \rightarrow M$, let $\gamma^- : [0, 1] \rightarrow M$ be the same curve traversed in the opposite orientation. For any oriented geodesic $\lambda \in \Lambda$ that meets γ transversely, let $\nu_\gamma^+(\lambda)$ be the number of positively oriented intersections of λ with γ . Define ν_γ^- similarly to be the number of negatively oriented intersections of λ with γ . Then, as Berwald observed, Crofton's formula remains valid in this case in the form

$$\mathcal{L}(\gamma) + \mathcal{L}(\gamma^-) = \int_{\Lambda} \nu_\gamma^+ \mu = \int_{\Lambda} \nu_\gamma^- \mu.$$

When Σ is symmetric, $\mathcal{L}(\gamma) = \mathcal{L}(\gamma^-)$ and the formula simplifies to

$$\mathcal{L}(\gamma) = \frac{1}{4} \int_{\Lambda} (\nu_\gamma^+ + \nu_\gamma^-) \mu.$$

This shows that, in the geodesically amenable case, the notion of L -length in M can be recovered from the knowledge of the space Λ of L -geodesics together with a measure μ on Λ .

This observation forms the basis of Pogorelov's solution of Hilbert's Fourth Problem in the symmetric case, as well as Szabó's more recent treatment. This also points out the importance of studying the path geometry defined by the L -geodesics. It is to this study that the next section is devoted.

2. PATH GEOMETRIES AND PROJECTIVE STRUCTURES

A Finsler structure on a surface M defines a 2-parameter family of oriented paths on M , one in every oriented direction through every point. This is a special case of the notion of *path geometry*. In this section, the basics of path geometries on surfaces are recalled and a notion of *generalized path geometry* is introduced so as to correspond with generalized Finsler structures. Cartan's solution of the equivalence problem for (generalized) path geometries is then reviewed with the purpose of developing an effective test for when a path geometry is locally equivalent to the 'flat' example of lines in the plane. Finally, this is applied to derive the classical conditions (due to Berwald) for a Finsler structure to have its geodesic path geometry be equivalent to that of the lines in the plane.

2.1. Path geometries and generalized path geometries. Roughly speaking, a *path geometry* on a surface M is a 2-parameter family Λ of curves on M with the property that for every point $x \in M$ and every line $L \subset T_x M$ there is a unique curve $\xi \in \Lambda$ with the property that ξ passes through x and has L as its tangent line at x . The fundamental example to keep in mind is the family of lines in the Euclidean plane. The actual definition to be given below will refine this intuitive picture by incorporating appropriate notions of smoothness and independence.

2.1.1. Path geometry on surfaces. For any surface M , let TM , as usual, denote the tangent bundle, and let $\pi : \mathbb{P}(TM) \rightarrow M$ denote the projectivized tangent

bundle, whose fiber over $x \in M$ is the projectivization of $T_x M$, i.e., the space of lines through $0_x \in T_x M$. Given any smooth, immersed curve $\gamma : (a, b) \rightarrow M$, there is a canonical lift $\gamma_1 : (a, b) \rightarrow \mathbb{P}(TM)$ defined by the rule $\gamma_1(t) = T_{\gamma(t)} \gamma((a, b))$.

It is easy to characterize the canonical lifts of immersed curves in M as curves in $\mathbb{P}(TM)$ in terms of local geometric structures on $\mathbb{P}(TM)$. In fact, as is well-known, $\mathbb{P}(TM)$ carries the structure of a 3-dimensional contact manifold. This contact structure is defined as follows: Since $\pi : \mathbb{P}(TM) \rightarrow M$ is a submersion, for each $L \in \mathbb{P}(TM)$, the linear map $\pi'(L) : T_L \mathbb{P}(TM) \rightarrow T_{\pi(L)} M$ is surjective. Hence, the inverse image, $E(L) = \pi'(L)^{-1}(L) \subset T_L \mathbb{P}(TM)$ is a canonically defined 2-plane in $T_L \mathbb{P}(TM)$. The 2-plane field E defines a contact structure on $\mathbb{P}(TM)$.

From the definition of the canonical lift γ_1 given above, it follows that γ_1 is tangent to E at all points, i.e., γ_1 is a contact curve. Moreover, since the projection $\gamma = \pi \circ \gamma_1$ is an immersion, it follows that γ_1 is also transverse to the fibers of π . Conversely, any contact curve $\phi : (a, b) \rightarrow \mathbb{P}(TM)$ which is transverse to the fibers of π is of the form $\phi = (\pi \circ \phi)_1$, and so is, in particular, a canonical lift.

Definition 2. A *path geometry* on a surface M is a foliation \mathcal{P} of $\mathbb{P}(TM)$ by contact curves, each of which is transverse to the fibers of $\pi : \mathbb{P}(TM) \rightarrow M$. A *local path geometry* on M is a foliation \mathcal{P} of an open subset $U \subset \mathbb{P}(TM)$ by contact curves, each of which is transverse to the fibers of π .

Note that, since a path geometry \mathcal{P} is a codimension 2 foliation, it makes sense to say that it defines a 2-parameter family of curves on the base space M , though this is somewhat imprecise. It could happen that the space of leaves of \mathcal{P} is non-Hausdorff.

In the case that there is a surface Λ and a submersion $\lambda : \mathbb{P}(TM) \rightarrow \Lambda$ whose fibers are the leaves of \mathcal{P} , the path geometry will be said to be *amenable*. This is the case, for example, for the path geometry that consists of the lines in the standard Euclidean plane. By contrast, the path geometry defined by the geodesics on a compact surface of constant negative curvature is very far from being amenable, as the geodesic flow is ergodic in this case.

2.1.2. Generalized path geometries. In this article, it will be necessary to work with path geometries that are only locally defined. In fact, it is useful to generalize the notion of a path geometry as follows:

Definition 3. A *generalized path geometry* on a 3-manifold Σ is a pair of transverse codimension 2 foliations $(\mathcal{P}, \mathcal{Q})$ with the property that the (unique) 2-plane field E that is tangent to both foliations defines a contact structure on Σ .

A path geometry \mathcal{P} on $\Sigma = \mathbb{P}(TM)$ in the classical sense is a special case of a generalized path geometry where the second foliation \mathcal{Q} is taken to be the fibers of the basepoint projection $\pi : \mathbb{P}(TM) \rightarrow M$.

It is convenient to introduce a notion of *orientability*. A generalized path geometry $(\Sigma, \mathcal{P}, \mathcal{Q})$ will be said to be *oriented* if a continuous choice of orientation of the leaves of each of \mathcal{P} and \mathcal{Q} has been made. A choice of orientation is then equivalent to the choice of two non-vanishing vector fields on Σ defined up to positive multiples, say, X_1 and X_3 , the first being tangent to the leaves of \mathcal{P} and the second being tangent to the leaves of \mathcal{Q} .

I skipped the index ‘2’ because I want to set $X_2 = [X_3, X_1]$ so as to correspond more closely to the notation established in the previous section. The hypothesis

that the 2-plane field E spanned by X_1 and X_3 be a contact plane field implies that the vector field X_2 is linearly independent from X_1 and X_3 .

Note that if I replace X_1 by $X_1^* = \lambda_1 X_1$ and X_3 by $X_3^* = \lambda_3 X_3$, where the functions λ_1 and λ_3 are both positive, then X_2 is replaced by

$$X_2^* = \lambda_1 \lambda_3 X_2 + d\lambda_1(X_3) X_1 - d\lambda_1(X_1) X_3.$$

This remark will be useful in the next subsection and farther along.

Finally, a 1-form α on Σ will be said to be \mathcal{P} -positive if it pulls back to each leaf of \mathcal{P} to be positive with respect to the specified orientation. The notion of \mathcal{Q} -positivity is defined similarly.

Example 2. Let \mathbb{R}^2 be endowed with its usual Euclidean metric and let $\Sigma \subset \mathbb{R}^2 \times \mathbb{R}^2$ be the set of pairs of points (\mathbf{x}, \mathbf{y}) in the plane satisfying $|\mathbf{x} - \mathbf{y}| = 1$. Let $\pi : \Sigma \rightarrow \mathbb{R}^2$ be the projection onto the second factor and let $\lambda : \Sigma \rightarrow \mathbb{R}^2$ be the projection onto the first factor. Let \mathcal{P} be the foliation defined by the fibers of λ and let \mathcal{Q} be the foliation defined by the fibers of π . Then it can be verified that $(\Sigma, \mathcal{P}, \mathcal{Q})$ is a generalized path geometry. Moreover, it can be oriented so that the clockwise orientation of the unit circles $\pi^{-1}(\mathbf{y})$ and $\lambda^{-1}(\mathbf{x})$ are each positive.

Note, by the way, that Σ does not induce a path geometry on \mathbb{R}^2 in the usual sense because through each point in the plane there pass *two* unit circles having the same (unoriented) tangent line.

2.1.3. Local realization. In spite of Example 2, every generalized path geometry $(\mathcal{P}, \mathcal{Q})$ on a 3-manifold Σ is locally identifiable with a local path geometry on a surface.

To show this, a (local) candidate surface M must be found. This can be done as follows: Let $\mathbf{u} \in \Sigma$ be chosen and let $U \subset \Sigma$ be an open neighborhood of \mathbf{u} on which the foliation \mathcal{Q} is amenable, i.e., so that there exists a smooth surface M and a smooth surjective submersion $\pi : U \rightarrow M$ so that the fibers of π are the leaves of \mathcal{Q} restricted to U . (Note that M and π are uniquely determined by U up to the natural notion of equivalence up to diffeomorphism.)

A canonical smooth map $\nu : \Sigma \rightarrow \mathbb{P}(TM)$ can now be defined as follows: For each $\mathbf{v} \in U$, let $\nu(\mathbf{v}) = \pi'(\mathbf{v})(T_{\mathbf{v}}\mathcal{P})$. Since the foliations \mathcal{Q} and \mathcal{P} are transverse and since the fibers of π are the leaves of \mathcal{Q} , it follows that $\pi'(\mathbf{v})(T_{\mathbf{v}}\mathcal{P})$ is a one-dimensional subspace of $T_{\pi(\mathbf{v})}M$, i.e., an element of $\mathbb{P}(T_{\pi(\mathbf{v})}M)$. Thus, ν is well-defined. The hypothesis that the plane field E be a contact field implies that ν is an immersion and hence, for dimension reasons, a local diffeomorphism. Shrinking U if necessary, it can be assumed that ν embeds U as an open subset of $\mathbb{P}(TM)$.

It is now not difficult to verify that ν is a contact diffeomorphism, i.e., it identifies the given contact structure on Σ with the one got by pulling back the canonical contact structure on $\mathbb{P}(TM)$ via ν .

In this way, the generalized path geometry on U is canonically identified with a local path geometry on M .

If the path geometry is oriented, there is an induced orientation on the surface M , defined as follows: Let X_1, X_2 and X_3 be chosen to correspond to the given orientation as in the last section. The vector fields X_1 and X_2 then span a 2-plane field H complimentary to the tangents to the fibers of π . Since the fibers of π are, by definition, connected, it follows that there is a unique orientation on the tangent spaces to M so that, for all $\mathbf{v} \in U$, the isomorphism $\pi'(\mathbf{v}) : H_{\mathbf{v}} \rightarrow T_{\pi(\mathbf{v})}M$

carries $(X_1(\mathbf{v}), X_2(\mathbf{v}))$ to an oriented basis of $T_{\pi(\mathbf{v})}M$. Because of the way X_2 was defined and because of the formula from §2.1.2 for how it changes when X_1 and X_3 are replaced by positive multiples, it follows that this orientation of M depends only on the orientation of the generalized path geometry, not on the specific choices of X_1 and X_3 .

2.1.4. Equivalence and duality. Two generalized path geometries $(\Sigma_1, \mathcal{P}_1, \mathcal{Q}_1)$ and $(\Sigma_2, \mathcal{P}_2, \mathcal{Q}_2)$ are said to be *equivalent* if there exists a diffeomorphism $\phi : \Sigma_1 \rightarrow \Sigma_2$ that satisfies $\phi(\mathcal{P}_1) = \mathcal{P}_2$ and $\phi(\mathcal{Q}_1) = \mathcal{Q}_2$.

The group of self-equivalences of a given generalized path geometry $(\Sigma, \mathcal{P}, \mathcal{Q})$ is known as its *automorphism group* or *symmetry group*. By a theorem of Cartan [Ca1], this symmetry group has the structure of a Lie group of dimension at most 8. In fact, if Σ is connected and \mathfrak{g} is the set of vector fields X on Σ having the property that the (locally defined) flow of X is a 1-parameter group of (local) symmetries of $(\Sigma, \mathcal{P}, \mathcal{Q})$, then \mathfrak{g} is a Lie algebra of dimension at most 8, with equality if and only if the generalized path geometry is locally equivalent to the one defined by the path geometry of straight lines in the plane. In this most symmetric case, the algebra \mathfrak{g} is isomorphic to $\mathfrak{sl}(3, \mathbb{R})$, the Lie algebra of infinitesimal projective transformations of the plane. (For further discussion of this point, see the next section.)

Finally, the reader may have noticed that there is a symmetry in the definition of a generalized path geometry, namely that the pair $(\mathcal{P}, \mathcal{Q})$ defines a generalized path geometry on Σ if and only if the pair $(\mathcal{Q}, \mathcal{P})$ also defines a generalized path geometry on Σ . This is known as the *dual* generalized path geometry. It is not generally true that a generalized path geometry is (locally) equivalent to its dual geometry.

2.2. The flat example: the projective plane. As a prelude to introducing Cartan's projective connection (a device for solving the equivalence problem for generalized path geometries) in the next section, I will now discuss the path geometry of the flat example.

2.2.1. Lines in the projective plane. Let V be \mathbb{R}^3 , thought of as the space of column vectors of height 3 with real entries, and let V^* be the space of row vectors of length 3 with real entries. The pairing $V^* \times V \rightarrow \mathbb{R}$ defined by matrix multiplication is non-degenerate and will be taken to be the canonical pairing (thus justifying the notation V^*).

Let $\mathbb{S} = (V \setminus \{0\})/\mathbb{R}^+$ denote the space of *oriented* lines through the origin in V . This space is diffeomorphic to the 2-sphere and canonically double covers $\mathbb{P}^2 = (V \setminus \{0\})/\mathbb{R}^*$, the space of (unoriented) lines through the origin in V . If $v \in V$ is any non-zero vector, let $[v]$ denote the corresponding point in \mathbb{S} . Note that, contrary to the usual usage, $[-v] \neq [v]$.

Let $\mathbb{S}^* = (V^* \setminus \{0\})/\mathbb{R}^+$ denote the 'dual' space of oriented lines through the origin in V^* , with $[\xi]$ denoting the oriented line corresponding to the non-zero element $\xi \in V^*$.

Let $\Sigma \subset \mathbb{S}^* \times \mathbb{S}$ denote the incidence correspondence

$$\Sigma = \{ ([\xi], [x]) \mid \xi \cdot x = 0 \}.$$

Then Σ is a smooth 3-manifold for which the two natural projections $\lambda : \Sigma \rightarrow \mathbb{S}^*$ and $\pi : \Sigma \rightarrow \mathbb{S}$ are smooth submersions. Moreover, Σ carries a canonical generalized

path geometry $(\mathcal{P}, \mathcal{Q})$ where the leaves of the foliation \mathcal{P} are the fibers of λ and the leaves of the foliation \mathcal{Q} are the fibers of π . The natural map $\nu : \Sigma \rightarrow \mathbb{P}(TS)$ is a 2-to-1 covering, with the points $([\pm\xi], [x])$ going to the same point under the map ν . Since the involution $([\xi], [x]) \mapsto ([-\xi], [x])$ maps the fibers of λ and π into fibers of these same maps, it follows that the generalized path geometry defined on Σ descends to a well-defined (classical) path geometry on \mathbb{S} . This is, of course, the classical geometry, where the paths on \mathbb{S} are the ‘great circles’.

The group $SL(3, \mathbb{R})$ acts on V on the left by matrix multiplication and on V^* on the right by matrix multiplication. This defines a left action on $V^* \times V$ by the rule

$$g \cdot (\xi, x) = (\xi g^{-1}, gx)$$

for $g \in SL(3, \mathbb{R})$. This action preserves the locus $\xi \cdot x = 0$ and descends to an action of $SL(3, \mathbb{R})$ on Σ which commutes via the projections λ and π with the natural actions of $SL(3, \mathbb{R})$ on \mathbb{S}^* and \mathbb{S} respectively. Each of these actions is transitive and effective (i.e., only the identity in $SL(3, \mathbb{R})$ acts as the identity transformation). In particular, the group $SL(3, \mathbb{R})$ acts as a group of symmetries of the induced path geometry on \mathbb{S} . It is a classical fact that $SL(3, \mathbb{R})$ is actually the full group of orientation preserving symmetries of this path geometry.

Let $x_0 = {}^t(1, 0, 0)$ and let $\xi_0 = (0, 0, 1)$. Let $P_1 \subset SL(3, \mathbb{R})$ be the stabilizer subgroup of $[x_0] \in \mathbb{S}$, let $P_2 \subset SL(3, \mathbb{R})$ be the stabilizer subgroup of $[\xi_0] \in \mathbb{S}^*$, and let $P = P_1 \cap P_2$. Note that P is the subgroup of upper triangular matrices in $SL(3, \mathbb{R})$ with all diagonal entries positive.

Now define a map $\tau : SL(3, \mathbb{R}) \rightarrow \Sigma$ by the rule

$$\tau(g) = g \cdot (\xi_0, x_0).$$

The map τ is surjective and its fibers are the left cosets of P , so that τ identifies $SL(3, \mathbb{R})/P$ with Σ .

Now let $\theta = g^{-1} dg$ be the canonical left-invariant 1-form on $SL(3, \mathbb{R})$, i.e., θ is the unique $\mathfrak{sl}(3, \mathbb{R})$ -valued 1-form on $SL(3, \mathbb{R})$ whose value at the identity $I_3 \in SL(3, \mathbb{R})$ is the identity mapping

$$\theta_{I_3} : T_{I_3} SL(3, \mathbb{R}) = \mathfrak{sl}(3, \mathbb{R}) \rightarrow \mathfrak{sl}(3, \mathbb{R}).$$

Since $\mathfrak{sl}(3, \mathbb{R})$ is the vector space of 3-by-3 matrices with trace zero, θ can be expanded in the form

$$\theta = \begin{pmatrix} \theta_0^0 & \theta_1^0 & \theta_2^0 \\ \theta_0^1 & \theta_1^1 & \theta_2^1 \\ \theta_0^2 & \theta_1^2 & \theta_2^2 \end{pmatrix}$$

where $\theta_0^0 + \theta_1^1 + \theta_2^2 = 0$, but the θ_b^a are otherwise linearly independent.

For comparison with the construction in the next section, I want to remark on the following properties of θ with respect to the oriented generalized path geometry on Σ :

First, there is the fact that $SL(3, \mathbb{R})$ acts on the left on all of the spaces $SL(3, \mathbb{R})$, Σ , \mathbb{S} , and \mathbb{S}^* in such a way that it commutes with the maps τ , π , and λ ; preserves the oriented generalized path geometry on Σ and the orientations and path geometries on \mathbb{S} and \mathbb{S}^* ; and leaves θ invariant.

Second, if $\sigma : \Sigma \rightarrow \mathrm{SL}(3, \mathbb{R})$ is any section of $\tau : \mathrm{SL}(3, \mathbb{R}) \rightarrow \Sigma$ and $\phi = \sigma^* \theta$, then the components ϕ_0^1 , ϕ_0^2 , and ϕ_1^2 have the following properties: First, the leaves of the foliation \mathcal{P} are the integral curves of $\phi_0^2 = \phi_1^2 = 0$ while the 1-form ϕ_0^1 is \mathcal{P} -positive. Second, the leaves of the foliation \mathcal{Q} are the integral curves of $\phi_0^1 = \phi_0^2 = 0$ while the 1-form ϕ_1^2 is \mathcal{Q} -positive.

Finally, note the Maurer-Cartan identity $d\theta = -\theta \wedge \theta$.

2.3. Cartan's projective connection. In [Ca1], Cartan introduced a device for determining when two generalized path geometries were equivalent, the so-called *projective connection*. This subsection contains an account of his results, modified slightly to take into account the orientations.

2.3.1. The bundle and its connection form. Recall that the subgroup $P \subset \mathrm{SL}(3, \mathbb{R})$ has been defined to be the subgroup consisting of the upper triangular matrices with all diagonal entries positive. Its Lie algebra is $\mathfrak{p} \subset \mathfrak{sl}(3, \mathbb{R})$, the subspace consisting of upper triangular matrices with trace zero.

Theorem 2 (Cartan). *Let $(\Sigma, \mathcal{P}, \mathcal{Q})$ be an oriented generalized path geometry. There exists a principal right P -bundle $\tau : B \rightarrow \Sigma$ and an $\mathfrak{sl}(3, \mathbb{R})$ -valued 1-form θ on B with the following properties:*

- (1) *For each $b \in B$, the map $\theta_b : T_b B \rightarrow \mathfrak{sl}(3, \mathbb{R})$ is an isomorphism and θ pulls back to each fiber of τ to be the canonical \mathfrak{p} -valued left-invariant 1-form;*
- (2) *$R_g^* \theta = g^{-1} \theta g$ for each $g \in P$;*
- (3) *For some (and hence any) section $\sigma : \Sigma \rightarrow B$, the pullback 1-form $\phi = \sigma^* \theta$ has the properties that, first, the leaves of the foliation \mathcal{P} are the integral curves of $\phi_0^2 = \phi_1^2 = 0$ while the 1-form ϕ_0^1 is \mathcal{P} -positive and, second, the leaves of the foliation \mathcal{Q} are the integral curves of $\phi_0^1 = \phi_0^2 = 0$ while the 1-form ϕ_1^2 is \mathcal{Q} -positive;*
- (4) *The curvature 2-form $\Theta = d\theta + \theta \wedge \theta$ satisfies*

$$\Theta = \begin{pmatrix} 0 & M \theta_0^1 \wedge \theta_0^2 & \Theta_2^0 \\ 0 & 0 & L \theta_1^2 \wedge \theta_0^2 \\ 0 & 0 & 0 \end{pmatrix}$$

for some functions L and M on B .

The pair (B, θ) is uniquely characterized by these four properties: If (B', θ') also satisfies them then there exists a unique bundle isomorphism $f : B \rightarrow B'$ covering the identity on Σ so that $f^* \theta' = \theta$. \square

For the proof, the reader can consult [Ca1] or the more modern treatment in [KN].

It should be noted that Cartan's proof is entirely constructive, i.e., he shows how to define B and θ by an algorithmic process that only involves differentiation and solving certain explicit linear equations.

In particular, suppose given a 3-manifold Σ on which there are defined two linearly independent vector fields X_1 and X_3 with the property that $X_2 = [X_3, X_1]$ is linearly independent from X_1 and X_3 . Consider the oriented generalized path geometry $(\Sigma, \mathcal{P}, \mathcal{Q})$ defined by the oriented integral curves of X_1 and X_3 . Then the bundle B and the 1-form θ for the oriented generalized path geometry $(\Sigma, \mathcal{P}, \mathcal{Q})$ can be constructed by a (somewhat involved) recipe from the knowledge of the iterated brackets of X_1 and X_3 . One does not need to be able to 'integrate' the vector fields X_1 or X_3 .

2.3.2. *Some applications.* From Theorem 2, Cartan draws several conclusions.

First, two oriented generalized path geometries $(\Sigma_1, \mathcal{P}_1, \mathcal{Q}_1)$ and $(\Sigma_2, \mathcal{P}_2, \mathcal{Q}_2)$ are equivalent if and only if there exists a diffeomorphism $f : B_1 \rightarrow B_2$ satisfying $f^*\theta_2 = \theta_1$.

Second, the group of symmetries of a given oriented generalized path geometry $(\Sigma, \mathcal{P}, \mathcal{Q})$ is isomorphic to the group of diffeomorphisms $f : B \rightarrow B$ that satisfy $f^*\theta = \theta$. Since the components of θ provide a coframing of B , it follows (see [Ko]) that this diffeomorphism group has the structure of a Lie group of dimension at most 8.

Third, for the flat example, $\Sigma \subset \mathbb{S}^* \times \mathbb{S}$, the pair (B, θ) can be taken to be $(\mathrm{SL}(3, \mathbb{R}), g^{-1} dg)$ while Σ is identified with $\mathrm{SL}(3, \mathbb{R})/P$. In particular, in this case, since $\Theta = d\theta + \theta \wedge \theta = 0$, it follows that the functions L and M vanish identically on the flat example.

Thus, if $(\Sigma, \mathcal{P}, \mathcal{Q})$ is to be locally equivalent to the flat example, the functions L and M must vanish identically. Conversely, if $L = M = 0$, then the Bianchi identity $d\Theta = \Theta \wedge \theta - \theta \wedge \Theta$ implies that Θ_0^0 vanishes identically as well, so that Θ itself vanishes identically, i.e., $d\theta = -\theta \wedge \theta$. It follows that if Σ is simply connected, then there exists a smooth mapping $f : B \rightarrow \mathrm{SL}(3, \mathbb{R})$ which satisfies $f(b \cdot g) = f(b) \cdot g$ for all $g \in P$ and which pulls back the canonical left-invariant form on $\mathrm{SL}(3, \mathbb{R})$ to θ on B . This mapping is a local diffeomorphism and is unique up to left translation by a constant in $\mathrm{SL}(3, \mathbb{R})$. In particular, this induces a local diffeomorphism $\bar{f} : \Sigma \rightarrow \mathrm{SL}(3, \mathbb{R})/P$ which is a local equivalence of oriented generalized path geometries. Thus, the vanishing of L and M is also a sufficient condition that $(\Sigma, \mathcal{P}, \mathcal{Q})$ be locally equivalent to the flat example.

Fourth, the Bianchi identity shows that L and M cannot be constant unless they are identically zero. Since any diffeomorphism $f : B \rightarrow B$ that fixes θ must also fix Θ and hence L and M , it follows that, if L or M is non-zero, then the group of symmetries of the generalized path geometry has dimension strictly less than 8.

2.3.3. *Computations.* Let $(\Sigma, \mathcal{P}, \mathcal{Q})$ be an oriented generalized path geometry with Cartan structure bundle and connection (B, θ) . If $\sigma : \Sigma \rightarrow B$ is any section, then $\phi = \sigma^*\theta$ is an $\mathfrak{sl}(3, \mathbb{R})$ -valued 1-form with four properties:

- (1) $\phi_0^1 \wedge \phi_0^2 \wedge \phi_0^2$ is a non-vanishing 3-form on Σ ;
- (2) the leaves of the foliation \mathcal{P} are the integral curves of $\phi_0^2 = \phi_1^2 = 0$ while the 1-form ϕ_0^1 pulls back to each such leaf to be a \mathcal{P} -positive 1-form;
- (3) the leaves of the foliation \mathcal{Q} are the integral curves of $\phi_0^1 = \phi_0^2 = 0$ while the 1-form ϕ_1^2 pulls back to each such leaf to be a \mathcal{Q} -positive 1-form;
- (4) the curvature 2-form $\Phi = d\phi + \phi \wedge \phi$ satisfies

$$\Phi = \begin{pmatrix} 0 & \overline{M} \phi_0^1 \wedge \phi_0^2 & \Phi_2^0 \\ 0 & 0 & \overline{L} \phi_1^2 \wedge \phi_0^2 \\ 0 & 0 & 0 \end{pmatrix}$$

for some functions \overline{L} and \overline{M} on Σ .

Conversely, if ϕ is any $\mathfrak{sl}(3, \mathbb{R})$ -valued 1-form on Σ with these four properties, then it is of the form $\sigma^*\theta$ for some section σ of B .

In particular, in order to show that a given oriented generalized path geometry is locally equivalent to the flat example, it suffices to construct a 1-form ϕ on Σ

satisfying the first three properties and

$$\Phi = d\phi + \phi \wedge \phi = \begin{pmatrix} 0 & 0 & \Phi_2^0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The Bianchi identity $d\Phi = \Phi \wedge \phi - \phi \wedge \Phi$ then implies that the component Φ_2^0 must also be zero.

2.4. The generalized path geometry of a Finsler structure. To each generalized Finsler structure (Σ, ω) there is canonically associated an oriented generalized path geometry $(\Sigma, \mathcal{P}, \mathcal{Q})$. The leaves of \mathcal{P} are the integral curves of \mathbf{X}_1 and the leaves of \mathcal{Q} are the integral curves of \mathbf{X}_3 . An orientation is fixed by declaring that ω_1 be \mathcal{P} -positive and ω_3 be \mathcal{Q} -positive.

2.4.1. The projective connection form. Keeping the structure equations and Bianchi identity notation as given in §1.3, consider the $\mathfrak{sl}(3, \mathbb{R})$ -valued 1-form

$$\phi = \begin{pmatrix} \frac{1}{3}(I\omega_3 - J\omega_2) & -K\omega_1 - \frac{1}{3}K_3\omega_2 & -\frac{2}{3}K_3\omega_1 - U\omega_2 + \frac{1}{3}(I_2 + J_3)\omega_3 \\ \omega_1 & \frac{1}{3}(I\omega_3 - J\omega_2) & -\frac{1}{3}(I_2 + J_3)\omega_2 - \omega_3 \\ \omega_2 & \omega_3 & -\frac{2}{3}(I\omega_3 - J\omega_2) \end{pmatrix}$$

where $U = K + \frac{1}{3}K_{33} + \frac{1}{3}K_3 I$. A short calculation yields

$$d\phi + \phi \wedge \phi = \begin{pmatrix} 0 & \overline{M}\omega_1 \wedge \omega_2 & \Phi_2^0 \\ 0 & 0 & \overline{L}\omega_3 \wedge \omega_2 \\ 0 & 0 & 0 \end{pmatrix},$$

where

$$\begin{aligned} 3\overline{M} &= -K_{31} + 3K_2, \\ 3\overline{L} &= -I_{23} - J_{33} - 2I(I_2 + J_3) - 6J. \end{aligned}$$

Since ϕ has all four properties listed in §2.3.3, it is the pullback of the projective connection form via a section of the projective structure bundle.

2.4.2. Projective flatness. The results of the last subsection coupled with the discussion in §§2.3.2-3 now combine to show that the generalized path geometry associated to a generalized Finsler structure is locally equivalent to that of the flat example if and only if its invariants satisfy

$$\begin{aligned} K_{31} - 3K_2 &= 0, \\ I_{23} + J_{33} + 2I(I_2 + J_3) + 6J &= 0. \end{aligned}$$

This characterization goes back at least to Rund [Ru, p. 261], but in some form is due to Berwald [Be], at least in the context of classical Finsler structures. (Since the characterization is local, the distinction between the classical and the generalized case is not important here.)

In the context of Hilbert's Fourth Problem, this gives a diffeomorphism invariant characterization of the Finsler Lagrangians whose geodesics can be mapped to the straight lines by some diffeomorphism. A generalized Finsler structure that satisfies these conditions will be said to be *projectively flat*.

3. PROJECTIVELY FLAT FINSLER STRUCTURES ON S^2 SATISFYING $K = 1$

The main object of this article is to classify the Finsler structures on S^2 that satisfy $K = 1$ and are projectively flat. However, it is useful to start in a more general setting, as it leads to a stronger result.

Thus, throughout this section, (Σ, ω) will be a connected, compact 3-manifold Σ with finite fundamental group endowed with a generalized Finsler structure ω satisfying the conditions that $K \equiv 1$ and that the induced generalized path geometry on Σ be projectively flat. The case where Σ is an actual Finsler structure on S^2 that satisfies $K \equiv 1$ and is projectively flat is an example, for such a Σ will be compact and its fundamental group will be \mathbb{Z}_2 .

It is not clear that the assumption that Σ have finite fundamental group is necessary for the conclusions of the theorems, as it is possible that compactness implies it. However, this assumption does simplify the discussion and is satisfied in the cases of interest, so it seems reasonable to include it.

It may seem that projective flatness plus constant curvature is so stringent an assumption that there will not be many such structures, even locally. However, surprisingly, it turns out that there are very many local examples, as was discovered by Funk [Fu1, Fu2]. In fact, the problem of characterizing the global solutions remained a concern throughout Funk's career [Fu3].

In this section, the first result will be Theorem 3, which asserts that any generalized Finsler structure (Σ, ω) on a simply connected, compact 3-manifold that is projectively flat and satisfies $K = 1$ is the double cover of a classical Finsler structure on \mathbb{S} with these properties. From this, the classification in the case of finite fundamental group follows.

The second result is a finiteness result, Theorem 4, which shows that, in a certain sense, two such structures that agree to sufficiently high order at corresponding points must be globally equivalent.

The final result is an existence result, Theorem 5, showing that non-Riemannian examples exist and depend on two essential parameters.

One surprising aspect of the proof is that it shows that the compact examples are naturally an open subset of an analytic family of examples, not all of which are compact.

3.1. The structure equations. Using the condition $K = 1$, the structure equations for the canonical coframing ω on Σ simplify to

$$(1) \quad \begin{aligned} d\omega_1 &= -\omega_2 \wedge \omega_3, \\ d\omega_2 &= (\omega_1 - I\omega_2) \wedge \omega_3, \\ d\omega_3 &= -(\omega_1 - J\omega_3) \wedge \omega_2, \end{aligned}$$

and the Bianchi identities reduce to:

$$(2) \quad \begin{aligned} dI &= J\omega_1 + I_2\omega_2 + I_3\omega_3, \\ dJ &= -I\omega_1 + J_2\omega_2 + J_3\omega_3. \end{aligned}$$

Since $K \equiv 1$, the first of the two conditions for projective flatness of the associated generalized path geometry, namely, $K_{31} - 3K_2 = 0$, is an identity, so the only remaining condition to impose is $I_{23} + J_{33} + 2I(I_2 + J_3) + 6J = 0$, so I assume this from now on.

The quantity $\frac{1}{3}(I_2 + J_3)$ turns up many places in what follows, so I will denote it by T . The condition for projective flatness then becomes

$$(3) \quad T_3 + 2IT + 2J = 0.$$

Now, the structure equations yield $d(I\omega_3 - J\omega_2) = 3T\omega_2\wedge\omega_3$. Applying the exterior derivative to both sides of this equation and using the structure equations again yields $0 = 3dT\wedge\omega_2\wedge\omega_3$, so it follows that $T_1 = 0$. Thus,

$$dT = T_2\omega_2 - 2(IT + J)\omega_3.$$

Differentiating this relation and wedging both sides with ω_2 yields

$$0 = (T_2\omega_1\wedge\omega_3 - 2(JT - I)\omega_1\wedge\omega_3)\wedge\omega_2,$$

so that $T_2 = 2(JT - I)$, i.e.,

$$(4) \quad dT = 2(JT - I)\omega_2 - 2(IT + J)\omega_3.$$

3.2. The canonical immersion. Now consider the $\mathfrak{sl}(3, \mathbb{R})$ -valued 1-form

$$(5) \quad \phi = \begin{pmatrix} \frac{1}{3}(I\omega_3 - J\omega_2) & -\omega_1 & -\omega_2 + T\omega_3 \\ \omega_1 & \frac{1}{3}(I\omega_3 - J\omega_2) & -\omega_3 - T\omega_2 \\ \omega_2 & \omega_3 & -\frac{2}{3}(I\omega_3 - J\omega_2) \end{pmatrix}.$$

By equations (1–4), the $\mathfrak{sl}(3, \mathbb{R})$ -valued 1-form ϕ satisfies $d\phi = -\phi\wedge\phi$.

Theorem 3. *Suppose that Σ is a compact, simply connected 3-manifold on which there is a generalized Finsler structure ω that satisfies $K = 1$ and is projectively flat. Define ϕ as above.*

- (1) *There exists a smooth immersion $g : \Sigma \rightarrow \mathrm{SL}(3, \mathbb{R})$, unique up to left translation by a constant matrix in $\mathrm{SL}(3, \mathbb{R})$, that satisfies $\phi = g^{-1}dg$.*
- (2) *The map g is a double cover onto its image $\bar{\Sigma} = g(\Sigma)$, a smoothly embedded 3-dimensional submanifold of $\mathrm{SL}(3, \mathbb{R})$.*
- (3) *The quotient generalized Finsler structure on $\bar{\Sigma} \simeq \Sigma/\mathbb{Z}_2$ is both amenable and geodesically amenable, with the projection $\bar{\Sigma} \rightarrow \mathbb{S} = \mathrm{SL}(3, \mathbb{R})/P_1$, resp. $\bar{\Sigma} \rightarrow \mathbb{S}^* = \mathrm{SL}(3, \mathbb{R})/P_2$, having the leaves of the induced foliation $\bar{\mathcal{Q}}$, resp. $\bar{\mathcal{P}}$, as fibers.*
- (4) *The induced canonical immersion $\bar{\nu} : \bar{\Sigma} \rightarrow T\mathbb{S}$ embeds $\bar{\Sigma}$ as a Finsler structure on \mathbb{S} whose geodesics are the standard lines in \mathbb{S} .*

Proof. Since Σ is simply connected and ϕ satisfies the Maurer-Cartan equation $d\phi + \phi\wedge\phi = 0$ (sometimes known as the ‘zero-curvature equation’), it follows by Cartan’s technique of the graph that there exists a smooth mapping $g : \Sigma \rightarrow \mathrm{SL}(3, \mathbb{R})$ so that $\phi = g^{-1}dg$ and that this mapping is unique up to left translation, i.e., if $\phi = h^{-1}dh$ for some other map $h : \Sigma \rightarrow \mathrm{SL}(3, \mathbb{R})$ then there exists a constant matrix $h_0 \in \mathrm{SL}(3, \mathbb{R})$ so that $h = h_0 g$.

Since the three subdiagonal entries of $g^{-1}dg = \phi$ form a coframing of Σ , it follows that g is an immersion.

Now, recall that $\mathbb{S} = \mathrm{SL}(3, \mathbb{R})/P_1$ where the mapping $[x] = \pi \circ \tau : \mathrm{SL}(3, \mathbb{R}) \rightarrow \mathbb{S}$ defined by $[x](g) = [g \cdot x_0]$ has the left cosets of P_1 as fibers. From the definition of ϕ , it then follows that $[x] \circ g : \Sigma \rightarrow \mathbb{S}$ is a submersion whose fibers are unions of integral curves of \mathbf{X}_3 , i.e., the leaves of \mathcal{Q} .

Since Σ is compact, $[x] \circ g$ must be a surjective submersion and each fiber of $[x] \circ g$ must be compact. In particular, each such fiber must be diffeomorphic to a disjoint union of a finite number of circles. Since each leaf of \mathcal{Q} must be a component of such a fiber, each leaf of \mathcal{Q} must be compact.

The submersion $[x] \circ g : \Sigma \rightarrow \mathbb{S}$ must therefore be a fibration. The exact homotopy sequence of this fibration includes the segment

$$\{\mathbf{1}\} \simeq \pi_1(\mathbb{S}) \longrightarrow \pi_0(F) \longrightarrow \pi_0(\Sigma) \simeq \{\mathbf{1}\},$$

where F is a typical fiber. Thus $\pi_0(F)$ must be trivial, i.e., the fibers of $[x] \circ g$ are connected, so that each fiber consists of precisely one leaf of \mathcal{Q} .

In sum, the generalized Finsler structure on Σ is amenable and I can regard the submersion $[x] \circ g : \Sigma \rightarrow \mathbb{S}$ as the canonical \mathcal{Q} -quotient mapping. (By the way, it is *not* generally true that the integral curves of \mathbf{X}_3 must be amenable for a generalized Finsler structure satisfying only $K = 1$, even when the underlying manifold Σ is compact with finite fundamental group. The hypothesis of projective flatness has been used in a crucial way in this argument.)

Let $\nu : \Sigma \rightarrow T\mathbb{S}$ be the induced canonical immersion as in §1.2. By Proposition 1, it follows that the image of each fiber $\Sigma_s = ([x] \circ g)^{-1}(s)$ in $T_s\mathbb{S}$ is a closed curve that is locally strictly convex towards 0_s for all $s \in \mathbb{S}$. The number of times that $\nu(\Sigma_s)$ winds around 0_s is independent of s . Moreover, this winding number must be 2 since the fundamental group of the unit circle bundle $U\mathbb{S} \subset T\mathbb{S}$ of any Riemannian metric g on \mathbb{S} is \mathbb{Z}_2 and the natural mapping $\beta : \Sigma \rightarrow U\mathbb{S}$ defined by $\beta(\mathbf{u}) = \nu(\mathbf{u})/|\nu(\mathbf{u})|_g$ is a covering covering map and hence must be 2-to-1.

Now consider the dual map $[\xi] \circ g : \Sigma \rightarrow \mathbb{S}^* = \mathrm{SL}(3, \mathbb{R})/P_2$. This mapping is also a submersion and its fibers are unions of integral curves of \mathbf{X}_1 . The same argument as before applies to show that $[\xi] \circ g$ must be a surjective submersion with each of its fibers being connected and hence consisting of a single, closed integral curve of \mathbf{X}_1 . Thus, the generalized Finsler structure on Σ is also geodesically amenable.

Now, the leaves of \mathcal{P} satisfy $\omega_2 = \omega_3 = 0$ and hence, from the definition of ϕ , it follows that each leaf of \mathcal{P} maps via g submersively onto a left coset of the circle subgroup

$$H = \left\{ \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \middle| \theta \in \mathbb{R} \right\}.$$

The image of such a left coset under the quotient map $[x] : \mathrm{SL}(3, \mathbb{R}) \rightarrow \mathbb{S}$ is a projective ‘line’, i.e., the set of oriented lines lying in a plane in V . In particular, it follows that the geodesics of the generalized Finsler structure on Σ map to the standard lines on \mathbb{S} .

Let $\Phi_t : \Sigma \rightarrow \Sigma$ be the time t flow of the vector field \mathbf{X}_1 . Since Σ is compact, this flow exists for all time. Since the structure equations show that $\mathcal{L}_{\mathbf{X}_1} \omega_1 = 0$, $\mathcal{L}_{\mathbf{X}_1} \omega_2 = \omega_3$, and $\mathcal{L}_{\mathbf{X}_1} \omega_3 = -\omega_2$, it follows that

$$\begin{aligned} \Phi_s^* \omega_1 &= \omega_1, \\ \Phi_s^* \omega_2 &= \cos s \omega_2 + \sin s \omega_3, \\ \Phi_s^* \omega_3 &= -\sin s \omega_2 + \cos s \omega_3. \end{aligned}$$

In particular, $\Phi_{2\pi} : \Sigma \rightarrow \Sigma$ satisfies $\Phi_{2\pi}^* \omega_i = \omega_i$, for $i = 1, 2, 3$. Consequently, $\Phi_{2\pi}^*$ must also preserve I , J , T , and hence ϕ . It is also evident that

$$g(\Phi_s(\mathbf{u})) = g(\mathbf{u}) \cdot \begin{pmatrix} \cos s & -\sin s & 0 \\ \sin s & \cos s & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so $g(\Phi_{2\pi}(\mathbf{u})) = g(\mathbf{u})$ for all $\mathbf{u} \in \Sigma$. In particular, $\nu \circ \Phi_{2\pi} = \nu$, so that $\Phi_{2\pi}$ acts as a deck transformation for the double cover $\beta : \Sigma \rightarrow U\mathbb{S}$ constructed above. It follows either that $\Phi_{2\pi}$ is the identity or else that it is a fixed point free involution.

Now $\Phi_{2\pi}$ cannot be the identity, since the image of a left coset of H is a line in \mathbb{S} whose tangential lift to $U\mathbb{S}$ is a generator of its fundamental group \mathbb{Z}_2 . Thus the double cover $\Sigma \rightarrow U\mathbb{S}$ must be a non-trivial double cover over each such lift, implying that the map g is 2-to-1 on each leaf of \mathcal{P} .

In sum, the map g commutes with the fixed point free involution $\Phi_{2\pi}$ and hence induces a smooth immersion of the quotient $\bar{\Sigma} = \Sigma/\mathbb{Z}_2$ into $\mathrm{SL}(3, \mathbb{R})$. This immersion must be an embedding since the induced map $\bar{\nu} : \bar{\Sigma} \rightarrow U\mathbb{S}$ is now seen to be a 1-sheeted covering and factors through g . \square

Corollary 1. *If (Σ, ω) is a generalized Finsler structure on a connected compact manifold Σ with finite fundamental group that satisfies $K \equiv 1$ and whose associated path geometry is locally flat, then either (Σ, ω) is a quotient of a projectively flat Finsler structure on \mathbb{S} by a finite subgroup of Finsler isometries or else it is a double cover of such a quotient.*

Proof. Let $\tilde{\Sigma} \rightarrow \Sigma$ be the simply connected cover and let $\tilde{\omega}$ be the pull-back of the generalized Finsler structure ω via the covering map. Since the deck transformations of this covering map preserve $\tilde{\omega}$, they must preserve the dual vector fields and, in particular, they must commute with the involution $\Phi_{2\pi} : \tilde{\Sigma} \rightarrow \tilde{\Sigma}$ constructed in the proof of the above theorem. There are now two cases to consider:

First, if $\Phi_{2\pi}$ is an element of the deck transformation group for $\tilde{\Sigma} \rightarrow \Sigma$, then it lies in the center of this group and so covers the identity on Σ . Dividing by the action of $\Phi_{2\pi}$ then induces a covering map $\bar{\Sigma} \rightarrow \Sigma$, so that Σ is a quotient of the Finsler structure $\bar{\Sigma}$ by the induced action of a finite group of Finsler isometries acting on \mathbb{S} .

Second, if $\Phi_{2\pi}$ is not an element of the deck transformation group, then because it commutes with the action of this group on $\tilde{\Sigma}$, it follows that there is a non-trivial involution $F : \Sigma \rightarrow \Sigma$ which preserves ω and makes the following diagram commute:

$$\begin{array}{ccc} \tilde{\Sigma} & \xrightarrow{\Phi_{2\pi}} & \tilde{\Sigma} \\ \downarrow & & \downarrow \\ \Sigma & \xrightarrow{F} & \Sigma \end{array}$$

Since F preserves ω and is not the identity, it cannot have any fixed points.

It follows that there exists a quotient $\Sigma \rightarrow \hat{\Sigma}$ whose deck transformation group is generated by the involution F . A simple diagram chase now shows that $\hat{\Sigma}$ is a quotient of $\bar{\Sigma} = \tilde{\Sigma}/\mathbb{Z}_2$. \square

3.3. The induced Riemannian metric on \mathbb{S}^* . In view of this theorem and its corollary, no essential generality is lost by assuming that (Σ, ω) is a Finsler structure on \mathbb{S} whose geodesics are the standard ones and that Σ has been embedded into $\text{SL}(3, \mathbb{R})$ via a mapping $g : \Sigma \rightarrow \text{SL}(3, \mathbb{R})$ that satisfies (5) so I will be assuming this from now on.

Since the integral curves of \mathbf{X}_1 are periodic of period 2π on Σ , it follows that the flow of \mathbf{X}_1 defines a free action of the unit circle on Σ whose orbits are the leaves of \mathcal{P} , i.e., the fibers of the submersion $[\xi] : \Sigma \rightarrow \mathbb{S}^*$. Thus, Σ can be regarded as a principal S^1 -bundle over \mathbb{S}^* .

The structure equations will now show that there are certain semi-basic tensors for this projection which are invariant under this S^1 -action and hence yield well-defined tensors on \mathbb{S}^* .

For example, since $\mathcal{L}_{\mathbf{X}_1} \omega_2 = \omega_3$ and $\mathcal{L}_{\mathbf{X}_1} \omega_3 = -\omega_2$, it follows that

$$\mathcal{L}_{\mathbf{X}_1} \omega_3 \wedge \omega_2 = 0 \quad \text{and} \quad \mathcal{L}_{\mathbf{X}_1} ((\omega_3)^2 + (\omega_2)^2) = 0.$$

In particular, there exists a 2-form dA on \mathbb{S}^* so that $[\xi]^*(dA) = \omega_3 \wedge \omega_2$ and there exists a Riemannian metric ds^2 on \mathbb{S}^* so that $[\xi]^*(ds^2) = (\omega_3)^2 + (\omega_2)^2$. From now on, I will consider \mathbb{S}^* to be an oriented Riemannian manifold with metric ds^2 and area form $dA > 0$.

Now the structure equations show that the map $\mu : \Sigma \rightarrow T\mathbb{S}^*$ defined by the formula $\mu(\mathbf{u}) = [\xi]'(\mathbf{u})(\mathbf{X}_3(\mathbf{u}))$ embeds Σ as the unit circle bundle of \mathbb{S}^* endowed with the metric ds^2 . In particular, the 1-forms ω_3 and ω_2 can be regarded as the tautological 1-forms on this circle bundle. The structure equations then show that the 1-form

$$\rho = -\omega_1 + I\omega_2 + J\omega_3$$

is the Levi-Civita connection for Σ regarded as this unit circle bundle. Note that this implies that the Gauss curvature K of this metric, which, by definition, satisfies $d\rho = -K\omega_3 \wedge \omega_2$, must be

$$K = 1 - I_3 + J_2 - I^2 - J^2.$$

3.4. A complex form of the structure equations. Since \mathbb{S}^* inherits a canonical metric and orientation from Σ , it follows that it has a unique structure as a Riemann surface for which the metric ds^2 is conformal and the area form dA is a positive $(1, 1)$ -form. A complex valued 1-form α on \mathbb{S}^* is of type $(1, 0)$ with respect to this complex structure if and only if it satisfies $[\xi]^*(\alpha) = A(\omega_3 + i\omega_2)$ for some function A on Σ .

For simplicity, set $\zeta = \omega_3 + i\omega_2$. Then the structure equations imply

$$(6) \quad d\zeta = -i\rho \wedge \zeta$$

where, as before, $\rho = -\omega_1 + I\omega_2 + J\omega_3$.

In this notation, equation §3.1.4 becomes

$$(7) \quad dT = -(T - i)(I + iJ)\zeta - (T + i)(I - iJ)\bar{\zeta}.$$

Taking the exterior derivative of both sides of this equation and using the structure equations then yields

$$0 = 2(J_2 - I_3 + 3(T^2 - I^2 - J^2))\omega_3 \wedge \omega_2,$$

so it follows that $I_3 - J_2 = 3(T^2 - I^2 - J^2)$. Using this relation, the equations (2) for dI and dJ can be written in the form

$$(8) \quad d((I + iJ)\zeta) = \frac{1}{2}(2I^2 + 2J^2 - 3T^2 - 3iT)\zeta \wedge \bar{\zeta}.$$

Note also that $K = 1 + 2I^2 + 2J^2 - 3T^2$, so that,

$$(9) \quad d\rho = -\frac{i}{2}(1 + 2I^2 + 2J^2 - 3T^2)\zeta \wedge \bar{\zeta}.$$

Conversely, the satisfaction of equations (6–9) is equivalent to the satisfaction of the system $d\phi + \phi \wedge \phi = 0$, once the translation of notation is made.

3.5. A holomorphic cubic differential. I now want to investigate the consequences of equations (6–9). To this end, it turns out to be useful to define $a = -(I + iJ)/(T + i)$. (Since T is real, $T + i$ can never vanish, so this formula does indeed define a smoothly.) In terms of a , the structure equations now take the simpler form

$$(10) \quad \begin{aligned} d\zeta &= -i\rho \wedge \zeta, \\ d\rho &= -\frac{i}{2}(1 + 2(T^2 + 1)|a|^2 - 3T^2)\zeta \wedge \bar{\zeta} \\ dT &= (T^2 + 1)(a\zeta + \bar{a}\bar{\zeta}) \\ d(a\zeta) &= \frac{3}{2}T\zeta \wedge \bar{\zeta}. \end{aligned}$$

Expanding and rearranging this last equation yields

$$(da - ia\rho + \frac{3}{2}T\bar{\zeta}) \wedge \zeta = 0.$$

It follows that there exists a complex valued function b on Σ so that

$$(11) \quad da = ia\rho + (b + a^2T)\zeta - \frac{3}{2}T\bar{\zeta}.$$

(Naming the ζ -coefficient $(b + a^2T)$ instead of b simplifies the next formula.) Differentiating (11) and using the structure equations yields

$$0 = (db - 2ib\rho + a\bar{\zeta}) \wedge \zeta.$$

Thus, there exists a complex valued function c on Σ so that

$$(12) \quad db = 2ib\rho + (c + 2abT - \frac{2}{9}a^3)\zeta - a\bar{\zeta}.$$

(Naming the ζ -coefficient in this equation $(c + 2abT - \frac{2}{9}a^3)$ instead of c simplifies the next formula). Finally, differentiating (12) and using the equations found so far yields the decisive formula

$$(13) \quad 0 = (dc - 3ic\rho) \wedge \zeta.$$

From this it follows that there is a cubic differential γ on \mathbb{S}^* satisfying $[\xi]^*(\gamma) = c\zeta^3$ and, moreover, that this differential form is holomorphic.

3.6. A vanishing theorem and its consequences. Now \mathbb{S}^* is topologically a 2-sphere, so by the Uniformization Theorem it must be equivalent as a Riemann surface to the Riemann sphere \mathbb{CP}^1 . Since there are no non-zero holomorphic differentials of positive degree on \mathbb{CP}^1 , it follows that $\gamma \equiv 0$, or equivalently, $c \equiv 0$. Thus, in complex form, the structure equations on Σ have finally been reduced to

$$(14) \quad \begin{aligned} d\zeta &= -i\rho\wedge\zeta, \\ d\rho &= -\frac{i}{2}(1+2(T^2+1)|a|^2-3T^2)\zeta\wedge\bar{\zeta}, \\ dT &= (T^2+1)(a\zeta+\bar{a}\bar{\zeta}), \\ da &= ia\rho+(b+a^2T)\zeta-\frac{3}{2}T\bar{\zeta}, \\ db &= 2ib\rho+(2abT-\frac{2}{9}a^3)\zeta-a\bar{\zeta}. \end{aligned}$$

Note that these formulae give expressions for the exterior derivative of every quantity appearing in the formulae. Moreover, exterior differentiation of these formulae simply yield identities, so that there are no further relations to be found by this method.

I will refer to the map $(T, a, b) : \Sigma \rightarrow \mathbb{R} \times \mathbb{C} \times \mathbb{C}$ as the *structure function* of the generalized Finsler structure ω .

In the next three subsections, I am going to prove a uniqueness theorem and an existence theorem. Since the constructions are somewhat long, I am going to state the results here, hoping that knowing the results will provide motivation to the reader for following the constructions.

Theorem 4. *Suppose that Σ and Σ^* are two connected, simply-connected and compact 3-manifolds endowed with generalized Finsler structures ω and ω^* , respectively, that are each projectively flat and satisfy $K = 1$. Let (T, a, b) and (T^*, a^*, b^*) respectively denote the real and complex-valued functions constructed from ω and ω^* by the formulae in §3.4–5. If there exist points $\mathbf{u} \in \Sigma$ and $\mathbf{u}^* \in \Sigma^*$ for which the equalities*

$$T(\mathbf{u}) = T^*(\mathbf{u}^*), \quad a(\mathbf{u}) = a^*(\mathbf{u}^*), \quad b(\mathbf{u}) = b^*(\mathbf{u}^*)$$

all hold, then there exists a unique diffeomorphism $f : \Sigma^ \rightarrow \Sigma$ so that $f^*(\omega) = \omega^*$ and $f(\mathbf{u}^*) = \mathbf{u}$.*

This uniqueness is accompanied by the following existence result:

Theorem 5. *There exists an open neighborhood U of $(0, 0, 0)$ in $\mathbb{R} \times \mathbb{C} \times \mathbb{C}$ so that for any $(T_0, a_0, b_0) \in U$ there exists a generalized Finsler structure ω on S^3 that is projectively flat, satisfies $K = 1$, and whose structure function (T, a, b) assumes the value (T_0, a_0, b_0) .*

Now, from the equations (14), the rank of the differential of the structure function can be computed in terms of the value of the structure function itself. Indeed, calculation shows that the structure function is an immersion except at the points where either $a = 0$ or else where there is a real number r so that $b = a^2r$ and $|a|^2 = \frac{9}{2}(1-3rT)/(1+9r^2)$. Since the generic point (T_0, a_0, b_0) in $\mathbb{R} \times \mathbb{C} \times \mathbb{C}$ does not satisfy either of these conditions, Theorem 5 implies that for the ‘generic’ generalized Finsler structures satisfying equations (14), the rank of the structure functions is 3 at most places. In other words, the images of the various structure functions are of dimension 3 at most places. This suggests that the images of these

structure functions should be the simultaneous level sets of a pair of functions defined on $\mathbb{R} \times \mathbb{C} \times \mathbb{C}$.

This is, indeed, the case. A little experimentation with equations (14) reveals that if new functions p and q are defined on Σ by the formulae

$$(15) \quad \begin{aligned} p &= \frac{|b|^2 + |a|^2 - \frac{1}{9}|a|^4 - \frac{9}{16}T^2 - \frac{27}{16}}{1 + T^2}, \\ q &= \frac{\frac{1}{3}(b\bar{a}^2 + \bar{b}a^2) + |a|^2T - \frac{9}{8}T}{1 + T^2}, \end{aligned}$$

then equations (14) imply

$$\begin{aligned} 3dp + \frac{2qdT}{1 + T^2} &= 0, \\ 3dq - \frac{2pdT}{1 + T^2} &= 0. \end{aligned}$$

Thus, the complex valued function w defined by

$$w = (p + iq)^3 \frac{(1 - iT)}{(1 + iT)}$$

must be constant on Σ . Especially, note that the function $W = p^2 + q^2$ is constant on Σ , as this remark will be important in what follows.

Note that w can be regarded as the result of composing the structure function $(T, a, b) : \Sigma \rightarrow \mathbb{R} \times \mathbb{C} \times \mathbb{C}$ with a \mathbb{C} -valued function $\mathbf{w} : \mathbb{R} \times \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$. Thus, the constancy of w can be regarded as the statement that the image of the structure function lies in a level surface of \mathbf{w} .

3.7. A differential ideal. In this section, I will construct a differential ideal on the manifold $X = \mathrm{SL}(3, \mathbb{R}) \times \mathbb{R} \times \mathbb{C} \times \mathbb{C}$ whose integral manifolds will correspond to the local and global solutions of the structure equations (14).

Let $\mathbf{g} : X \rightarrow \mathrm{SL}(3, \mathbb{R})$, $\mathbf{T} : X \rightarrow \mathbb{R}$, $\mathbf{a} : X \rightarrow \mathbb{C}$, and $\mathbf{b} : X \rightarrow \mathbb{C}$ denote the projections onto the first, second, third, and fourth factors of X respectively. Set

$$\mathbf{g}^{-1} d\mathbf{g} = \begin{pmatrix} \phi_1^1 & \phi_2^1 & \phi_3^1 \\ \phi_1^2 & \phi_2^2 & \phi_3^2 \\ \phi_1^3 & \phi_2^3 & \phi_3^3 \end{pmatrix}.$$

Then $d\phi_j^i = -\phi_k^i \wedge \phi_j^k$ and $\phi_1^1 + \phi_2^2 + \phi_3^3 = 0$, but this is the only linear relation among these nine 1-forms.

Define \mathbf{I} and \mathbf{J} on X by $\mathbf{I} + i\mathbf{J} = -(\mathbf{T} + i)\mathbf{a}$, set $\zeta = \phi_2^3 + i\phi_1^3$, and set $\rho = -\phi_1^2 + \mathbf{I}\phi_1^3 + \mathbf{J}\phi_2^3$.

Now consider the eleven 1-forms $\theta_0, \dots, \theta_{10}$ on X defined by the equations

$$\begin{aligned}\theta_0 &= \phi_1^1 - \frac{1}{3}(\mathbf{I}\phi_2^3 - \mathbf{J}\phi_1^3) \\ \theta_1 &= \phi_2^2 - \frac{1}{3}(\mathbf{I}\phi_2^3 - \mathbf{J}\phi_1^3) \\ \theta_2 &= \phi_3^3 + \frac{2}{3}(\mathbf{I}\phi_2^3 - \mathbf{J}\phi_1^3) \\ \theta_3 &= \phi_2^1 + \phi_1^2 \\ \theta_4 &= \phi_3^1 + \phi_1^3 - \mathbf{T}\phi_2^3 \\ \theta_5 &= \phi_3^2 + \phi_2^3 + \mathbf{T}\phi_1^3 \\ \theta_6 &= d\mathbf{T} - (\mathbf{T}^2 + 1)(\mathbf{a}\zeta + \bar{\mathbf{a}}\bar{\zeta}) \\ \theta_7 + i\theta_8 &= d\mathbf{a} - i\mathbf{a}\rho - (\mathbf{b} + \mathbf{a}^2\mathbf{T})\zeta + \frac{3}{2}\mathbf{T}\bar{\zeta} \\ \theta_9 + i\theta_{10} &= d\mathbf{b} - 2i\mathbf{b}\rho - (2\mathbf{a}\mathbf{b}\mathbf{T} - \frac{2}{9}\mathbf{a}^3)\zeta + \mathbf{a}\bar{\zeta}.\end{aligned}$$

These forms satisfy the relation $\theta_0 + \theta_1 + \theta_2 = 0$ but are otherwise linearly independent. Thus, they generate a Pfaffian system Θ of rank 10 on X .

The interest in this Pfaffian system is explained by the following result.

Theorem 6. *Let Σ be a compact, simply-connected 3-manifold and let ω be a generalized Finsler structure on Σ that is projectively flat and satisfies $K = 1$. Let $g : \Sigma \rightarrow \text{SL}(3, \mathbb{R})$ be its canonical immersion, and let $(T, a, b) : \Sigma \rightarrow \mathbb{R} \times \mathbb{C} \times \mathbb{C}$ be its structure function.*

Then the map $(g, T, a, b) : \Sigma \rightarrow X$ immerses Σ as an integral manifold of Θ and is a double cover onto its image, which is diffeomorphic to $\text{SO}(3)$ and which defines a Finsler structure on \mathbb{S} .

Conversely, if $N \subset X$ is any 3-dimensional integral manifold of Θ , then the triple $\omega = (\phi_1^2, \phi_1^3, \phi_2^3)$ defines a generalized Finsler structure on N that is projectively flat and satisfies $K = 1$. Moreover, if N is compact, then N is diffeomorphic to $\text{SO}(3)$ and the pair (N, ω) is diffeomorphic to a Finsler structure on \mathbb{S} endowed with its canonical coframing.

Proof. Essentially all the pieces of the proof have been assembled in the previous sections. That the map (g, T, a, b) immerses Σ as an integral manifold of Θ is a consequence of the way that g was defined to satisfy $g^{-1}dg = \phi$ in §3.2 and the vanishing theorem of §3.6, which established the equations (14). The combination of these results shows that $(g, T, a, b)^*(\theta_j) = 0$ for $0 \leq j \leq 10$. Moreover, by the theorem of §3.2, it follows that the map g double covers onto its image, which defines a Finsler structure on \mathbb{S} . Since the structure function (T, a, b) is defined in terms of the coframing ω , it must be invariant under the involution that defines the non-trivial deck transformation of this cover, so the full map (g, T, a, b) must also be a double cover onto its image.

Conversely, since the dimension of X is 13 and the rank of the Pfaffian system Θ is 10, any integral manifold $\Sigma \subset X$ of Θ cannot have dimension more than 3. Moreover, since the three 1-forms ϕ_1^2 , ϕ_1^3 , and ϕ_2^3 are linearly independent modulo Θ , it follows that, on any 3-dimensional integral manifold N of Θ , these three 1-forms must define a coframing. The vanishing of the forms $\theta_0, \dots, \theta_5$ on such an N imply that ϕ has the form of equation (5) on N . Of course this implies that if one defines $\omega_1 = \phi_1^2$, $\omega_2 = \phi_1^3$, and $\omega_3 = \phi_2^3$, then $\omega = (\omega_1, \omega_2, \omega_3)$ will satisfy the structure equations of a generalized Finsler structure on N that is projectively flat and satisfies $K = 1$.

Finally, since the projection $g : N \rightarrow SL(3, \mathbb{R})$ immerses N in such a way that $g^{-1}dg$ on N has the form (5), the arguments of §3.2 show that, if it is compact, it must be derived from a Finsler structure on \mathbb{S} and hence must have the stated topology and geometry. \square

3.8. The uniqueness theorem. The materials are now assembled to prove the uniqueness theorem of §3.6. The crucial observation is that, because Θ is a Pfaffian system of rank 10 on the 13-manifold X , any two integral manifolds of dimension 3 that intersect must be equal in a neighborhood of any point of intersection.

Proof of Theorem 4. Suppose that (Σ, ω) and (Σ^*, ω^*) satisfy the hypotheses of the theorem and that $\mathbf{u} \in \Sigma$ and $\mathbf{u}^* \in \Sigma^*$ satisfy $(T, a, b)(\mathbf{u}) = (T^*, a^*, b^*)(\mathbf{u}^*)$. Since the canonical immersions $g : \Sigma \rightarrow SL(3, \mathbb{R})$ and $g^* : \Sigma^* \rightarrow SL(3, \mathbb{R})$ are only well-defined up to left translation by constants, they can be uniquely specified by requiring that $g(\mathbf{u}) = g^*(\mathbf{u}) = I_3$.

The theorem of §3.7 now implies that each of the maps $(g, T, a, b) : \Sigma \rightarrow X$ and $(g^*, T^*, a^*, b^*) : \Sigma^* \rightarrow X$ is a double cover onto a compact 3-dimensional integral manifold of Θ . The hypothesis that the structure maps at \mathbf{u} and \mathbf{u}^* assume the same value implies that these two image integral manifolds intersect in at least one point. Thus, they must be equal, say, $(g, T, a, b)(\Sigma) = (g^*, T^*, a^*, b^*)(\Sigma^*) = N$. It follows that there is a unique diffeomorphism $f : \Sigma^* \rightarrow \Sigma$ so that $f(\mathbf{u}^*) = \mathbf{u}$ and $(g^*, T^*, a^*, b^*) = (g, T, a, b) \circ f$. Since, by construction, the generalized Finsler structures ω and ω^* are pulled back from the canonical one induced on the integral manifold N , it follows that $f^*(\omega) = \omega^*$, as desired. \square

3.9. The existence theorem. It remains to determine ‘how many’ compact integral manifolds of Θ there are.

There are some: The codimension 5 submanifold $Z \subset X$ defined by the equations $\mathbf{T} = \mathbf{a} = \mathbf{b} = 0$ is an integral manifold of the five 1-forms $\theta_6, \dots, \theta_{10}$. On Z , the remaining 1-forms in Θ reduce to the system $\{\phi_1^1, \phi_2^2, \phi_3^3, \phi_2^1 + \phi_1^2, \phi_2^3 + \phi_3^2, \phi_1^3 + \phi_3^1\}$. The integral manifolds of this latter system are the left cosets of $SO(3) \subset SL(3, \mathbb{R})$. By construction, these integral manifolds correspond to Finsler structures on \mathbb{S} that satisfy $I = J = 0$, i.e., Riemannian metrics of Gauss curvature 1 on \mathbb{S} whose geodesics are the standard straight lines.

The Riemannian solutions may be regarded as the trivial ones. I will now show that there are non-trivial ones that are, in a sense, close to Riemannian. The first step is the following proposition, whose proof is a straightforward calculation from the structure equations, so I omit it.

Proposition 3. *The Pfaffian system Θ is Frobenius. In particular, X has a foliation \mathcal{F} of codimension 10 whose leaves are maximal integral manifolds of Θ .* \square

Proof of Theorem 5. Consider the functions \mathbf{p} and \mathbf{q} defined on $\mathbb{R} \times \mathbb{C} \times \mathbb{C}$ by the formulae

$$\mathbf{p} = \frac{|\mathbf{b}|^2 + |\mathbf{a}|^2 - \frac{1}{9}|\mathbf{a}|^4 - \frac{9}{16}\mathbf{T}^2 - \frac{27}{16}}{1 + \mathbf{T}^2},$$

$$\mathbf{q} = \frac{\frac{1}{3}(\mathbf{b}\bar{\mathbf{a}}^2 + \bar{\mathbf{b}}\mathbf{a}^2) + |\mathbf{a}|^2\mathbf{T} - \frac{9}{8}\mathbf{T}}{1 + \mathbf{T}^2}.$$

A calculation shows that the 1-forms θ_{11} and θ_{12} defined by the formulae

$$\theta_{11} = 3d\mathbf{p} + \frac{2\mathbf{q}d\mathbf{T}}{1 + \mathbf{T}^2} \quad \text{and} \quad \theta_{12} = 3d\mathbf{q} - \frac{2\mathbf{p}d\mathbf{T}}{1 + \mathbf{T}^2}$$

are linear combinations of the forms $\theta_6, \dots, \theta_{10}$. It follows that the complex valued function \mathbf{w} defined by

$$\mathbf{w} = (\mathbf{p} + i\mathbf{q})^3 \frac{(1 - i\mathbf{T})}{(1 + i\mathbf{T})}$$

must be constant on each leaf of \mathcal{F} .

In particular, the function $\mathbf{W} = \mathbf{p}^2 + \mathbf{q}^2$ is constant on each leaf of \mathcal{F} . Now the Taylor series expansion of \mathbf{W} in a neighborhood of $(\mathbf{T}, \mathbf{a}, \mathbf{b}) = (0, 0, 0) = \mathbf{0}$ is of the form

$$\mathbf{W} = \frac{27}{8} \left(\frac{27}{8} - \frac{3}{4}\mathbf{T}^2 - |\mathbf{a}|^2 - |\mathbf{b}|^2 \right) + \dots$$

where the omitted terms vanish to order at least 3 at $\mathbf{0}$. It follows that $\mathbf{0}$ is a non-degenerate local maximum of \mathbf{W} on $\mathbb{R} \times \mathbb{C} \times \mathbb{C}$. In particular, there exists a compact domain $D \subset \mathbb{R} \times \mathbb{C} \times \mathbb{C}$ containing $\mathbf{0}$ in its interior and whose smooth boundary is a compact component of a level set of \mathbf{W} . Because \mathbf{W} is constant on the leaves of \mathcal{F} , it follows that any leaf of \mathcal{F} that intersects $\mathrm{SL}(3, \mathbb{R}) \times D \subset X$ must lie entirely inside $\mathrm{SL}(3, \mathbb{R}) \times D$.

Consider the natural projection $h : \mathrm{SL}(3, \mathbb{R}) \times D \rightarrow \mathrm{SL}(3, \mathbb{R})/P$. (Recall that P is the subgroup of upper triangular matrices with positive entries on the diagonal.) The leaves of \mathcal{F} are transverse to the fibers of h , so h restricts to each leaf L of \mathcal{F} to be a local diffeomorphism. In fact, the manifest invariance of Θ under the natural left action of $\mathrm{SL}(3, \mathbb{R})$ on X combined with the compactness of D shows that, for any leaf L of \mathcal{F} that lies in $\mathrm{SL}(3, \mathbb{R}) \times D$, every point of $\mathrm{SL}(3, \mathbb{R})/P \simeq \mathrm{SO}(3)$ is evenly covered by the local diffeomorphism $h : L \rightarrow \mathrm{SL}(3, \mathbb{R})/P$. Thus, each such L must be a covering space of $\mathrm{SO}(3)$ and hence must be compact. By the previous theorem, such a compact leaf must be diffeomorphic to $\mathrm{SO}(3)$ and hence the map h restricts to each leaf L to be a diffeomorphism onto $\mathrm{SL}(3, \mathbb{R})/P$.

For any $(T_0, a_0, b_0) \in D$, a generalized Finsler structure ω that is projectively flat and satisfies $K = 1$ and whose structure function assumes the value (T_0, a_0, b_0) can now be constructed on S^3 by simply taking the universal cover of the leaf L of \mathcal{F} that passes through $(I_3, T_0, a_0, b_0) \in \mathrm{SL}(3, \mathbb{R}) \times D$. \square

Remark. Not all of the leaves of \mathcal{F} on X are compact. The topology of the non-compact leaves is yet to be determined, but for some information, see §4.

4. CONNECTIONS WITH THE TREATMENT OF FUNK

In this final section, I want to explain the relationship of the preceding sections with the earlier results of Funk [Fu2, Fu3] on Finsler metrics on the plane whose geodesics are the straight lines and whose curvature satisfies $K = 1$.

In fact, Funk found a complete local classification, though this was not apparent to me at first, as I found his results hard to understand and his arguments hard to follow. In this section, I will give a discussion of his results in language that will ease comparison with the other sections of this paper and then go on to use these results to write down explicit global examples.

4.1. Funk's results. One of the difficulties of reading Funk's work for global implications is that he works on the affine plane \mathbb{R}^2 rather than on the natural global object \mathbb{RP}^2 or its double cover, which I identify as \mathbb{S} as in §2.2. Since I will be working on \mathbb{S} (both locally and globally), his results require some translation before they can be compared with those of the previous sections.

4.1.1. Projectively parametrized lines in \mathbb{S} . Recall the notation of §2.2, where V is identified with \mathbb{R}^3 and \mathbb{S} (diffeomorphic to S^2) is the space of oriented lines through the origin in V . I will fix the standard volume form on \mathbb{R}^3 , i.e., the standard identification of $\Lambda^3(V)$ with \mathbb{R} . Thus, for any three vectors v_0, v_1 , and v_2 in V , the wedge product $v_0 \wedge v_1 \wedge v_2$ will be treated as a number and this identification is invariant under the natural action of $\mathrm{SL}(3, \mathbb{R})$ on V .

Given an oriented 2-dimensional subspace $E \subset V$, let $\mathbf{v} = (v_0, v_1)$ be an oriented basis of E . Then the oriented line $[E]$ in \mathbb{S} is defined as the oriented curve parametrized by the map $\gamma_{\mathbf{v}} : S^1 \rightarrow \mathbb{S}$ defined by the formula

$$\gamma_{\mathbf{v}}(s) = [\cos s \ v_0 + \sin s \ v_1]$$

together with the convention that S^1 be oriented so that ds is a positive 1-form on S^1 . The choice of oriented basis \mathbf{v} of E affects this parametrization but any two such choices will yield the same image line with the same orientation.

A choice of oriented basis \mathbf{v} also defines a linear functional $\alpha_{\mathbf{v}} : V \rightarrow \mathbb{R}$ by the rule $\alpha_{\mathbf{v}}(v) = v_0 \wedge v_1 \wedge v$. The resulting oriented line $[\alpha_{\mathbf{v}}] \in \mathbb{S}^*$ depends only on the oriented plane E , not on the choice of oriented basis \mathbf{v} , so I will simply identify the points of \mathbb{S}^* with the space of oriented lines in \mathbb{S} via the identification $[E] = [\alpha_{\mathbf{v}}]$.

4.1.2. Projectively parametrized geodesics. An open domain $\mathbb{D} \subset \mathbb{S}$ will be said to be *convex* if its intersection with each line in \mathbb{S} is connected. If \mathbb{D} is convex, then I will let $\mathbb{D}^* \subset \mathbb{S}^*$ denote the set of oriented lines in \mathbb{S} whose intersection with \mathbb{D} is non-empty. (Note that it is *not* generally true that $\mathbb{D}_1^* = \mathbb{D}_2^*$ implies that $\mathbb{D}_1 = \mathbb{D}_2$ unless $\mathbb{D}_1^* = \mathbb{S}^*$.)

A Finsler structure $\Sigma_{\mathbb{D}} \subset T\mathbb{D}$ will be said to have *linear geodesics* if each of its (oriented) geodesics is of the form $[E] \cap \mathbb{D}$ for some (unique) oriented 2-plane E .

I can now state one of Funk's results:

Theorem 7 (Funk). *Let $\mathbb{D} \subset \mathbb{S}$ be a convex domain in \mathbb{S} and suppose that there is a Finsler structure $\Sigma_{\mathbb{D}}$ on \mathbb{D} with linear geodesics and whose curvature satisfies $K = 1$. Then, for every oriented line $[E] \in \mathbb{D}^*$, there exists an oriented basis $\mathbf{v} = (v_0, v_1)$ of $[E]$ so that the parametrization $\gamma_{\mathbf{v}}$ has unit speed (i.e., is a $\Sigma_{\mathbb{D}}$ -curve).*

Remark. Funk's proof is based on the characterization of the Finsler-Gauss curvature as the 0-th order term in the self-adjoint form of the Jacobi equation for variation of geodesics. It is straightforward, so I will not reproduce it here, see [Fu2]. \square

4.1.3. The space of oriented metric lines. Theorem 7 suggests looking at the problem of determining when two oriented bases of a 2-plane E induce the same metric and orientation on the line $[E]$, and this is what I will do next.

Suppose that $\mathbf{v} = (v_0, v_1)$ and $\mathbf{w} = (w_0, w_1)$ be two oriented bases of E with the property that the two parametrizations $\gamma_{\mathbf{v}}$ and $\gamma_{\mathbf{w}}$ induce the same metric (and orientation) on the line $[E]$. Then there is a constant s_0 so that $\gamma_{\mathbf{v}}(s) = \gamma_{\mathbf{w}}(s-s_0)$. I.e., for all s ,

$$[\cos s \ v_0 + \sin s \ v_1] = [\cos(s-s_0) \ w_0 + \sin(s-s_0) \ w_1],$$

which can only hold if there exists a positive real number r so that, for all s ,

$$(\cos s \ v_0 + \sin s \ v_1) = r (\cos(s-s_0) \ w_0 + \sin(s-s_0) \ w_1).$$

This, in turn, is equivalent to

$$v_0 + i v_1 = r e^{is_0} (w_0 + i w_1).$$

In particular, the points $\llbracket v_0 + i v_1 \rrbracket$ and $\llbracket w_0 + i w_1 \rrbracket$ in $\mathbb{CP}^2 = \mathbb{P}(V \otimes \mathbb{C})$ are equal.

Now, let $\mathbb{RP}^2 \subset \mathbb{CP}^2$ denote the set of real points. Given any point z in $\mathbb{CP}^2 \setminus \mathbb{RP}^2$, it can be represented in the form $z = \llbracket v_0 + i v_1 \rrbracket$ for some linearly independent (real) vectors v_0 and v_1 in V . The plane E_z spanned by the pair $\mathbf{v} = (v_0, v_1)$ and the orientation for which this is an oriented basis are independent of the specific choice of v_0 and v_1 satisfying $z = \llbracket v_0 + i v_1 \rrbracket$. Likewise, the metric on $[E_z]$ for which $\gamma_{\mathbf{v}}$ is a unit speed parametrization does not depend on this choice, but only on z .

In this way, the open 4-manifold $\mathbb{CP}^2 \setminus \mathbb{RP}^2$ can be regarded as a space of oriented Riemannian metrics on lines in \mathbb{S} . The map $\tau : \mathbb{CP}^2 \setminus \mathbb{RP}^2 \rightarrow \mathbb{S}^*$ defined by $\tau(z) = [\alpha_{\mathbf{v}}]$ is a smooth submersion whose fiber over $[E] \in \mathbb{S}^*$ consists of a two-parameter family of Riemannian metrics on the oriented line $[E]$.

Theorem 7 can now be reformulated:

Theorem 7'. *Let $\mathbb{D} \subset \mathbb{S}$ be a convex domain in \mathbb{S} and suppose that there is a Finsler structure $\Sigma_{\mathbb{D}}$ on \mathbb{D} with linear geodesics and whose curvature satisfies $K = 1$. Then there exists a unique section $\sigma : \mathbb{D}^* \rightarrow \mathbb{CP}^2 \setminus \mathbb{RP}^2$ of the bundle $\tau : \mathbb{CP}^2 \setminus \mathbb{RP}^2 \rightarrow \mathbb{S}^*$ so that, for each $[E] \in \mathbb{D}^*$, the metric $\sigma([E])$ on $[E]$ agrees with the metric induced on $[E] \cap \mathbb{D}$ by $\Sigma_{\mathbb{D}}$. \square*

Remark. The section σ will be called the *canonical section* associated to $\Sigma_{\mathbb{D}}$.

4.1.4. Local Finsler structures and complex curves. The advantage of recasting Funk's result in this language is that it simplifies the statement of Funk's local characterization of these Finsler structures:

Theorem 8. *Let $\mathbb{D} \subset \mathbb{S}$ be a convex domain in \mathbb{S} and suppose that there is a Finsler structure $\Sigma_{\mathbb{D}}$ on \mathbb{D} with linear geodesics and whose curvature satisfies $K = 1$. Then the image of the canonical section $\sigma : \mathbb{D}^* \rightarrow \mathbb{CP}^2 \setminus \mathbb{RP}^2$ is a complex curve in \mathbb{CP}^2 .*

Conversely, if $C \subset \mathbb{CP}^2 \setminus \mathbb{RP}^2$ is a complex curve with the property that the map $\tau : C \rightarrow \mathbb{S}^$ is a diffeomorphism onto its image, then there exists a generalized Finsler structure $\Sigma_C \subset T\mathbb{S}$ with the property that $\tau(z)$ endowed with the metric z is a Σ_C -curve. Furthermore, this generalized Finsler structure has the lines in \mathbb{S} as geodesics and satisfies $K = 1$. \square*

Remarks. Theorem 8 is only partly equivalent to the corresponding results of [Fu2]. Indeed, it took me quite some time to make the translation of his results into this form. What Funk does have is a formula for the local Finsler structures with linear geodesics and curvature equal to 1 in terms of an arbitrary holomorphic function of one variable, which is locally the same as a complex curve in \mathbb{C}^2 .

This should be compared with the constructions of §§3.4–5, where manipulation of the structure equations of a projectively flat generalized Finsler structure with $K = 1$ leads to the introduction of a complex structure on the space of geodesics and then the discovery of a holomorphic cubic differential, in terms of which the local geometry can be recovered.

The main advantage of this version of Funk's local characterization theorem is that it is fully invariant under the action of the projective group $\mathrm{SL}(3, \mathbb{R})$, which Funk's description was not. Even in the later paper [Fu3], where Funk was concerned with a 'global' characterization of the Riemannian metrics on the plane

having constant curvature and linear geodesics among the Finsler structures on the plane with these properties, he is seriously hampered by not having a formulation that is invariant under the full projective group.

4.1.5. A global classification. Combining these results of Funk leads immediately to a characterization of the Finsler structures defined on the entire sphere \mathbb{S} with linear geodesics and with $K = 1$:

Theorem 9. *Let Σ be a Finsler structure on \mathbb{S} with linear geodesics that satisfies $K = 1$. The image of the associated canonical section $\sigma : \mathbb{S}^* \rightarrow \mathbb{CP}^2 \setminus \mathbb{RP}^2$ is then a smooth conic C (i.e., smooth algebraic curve of degree 2) without real points.*

Conversely, if $C \subset \mathbb{CP}^2$ is a smooth conic without real points, then it is the image of the canonical section of a unique Finsler structure Σ_C on \mathbb{S} with linear geodesics that satisfies $K = 1$.

Proof. By Theorem 8, the image C is a complex curve in $\mathbb{CP}^2 \setminus \mathbb{RP}^2$. Since it is the image of a section of a bundle whose base is diffeomorphic to a 2-sphere, it follows that C must be diffeomorphic to a 2-sphere, i.e., it must be a smoothly embedded rational curve in \mathbb{CP}^2 without real points. Now, it is a standard result in algebraic geometry [GH, Chapter 1] that any smoothly embedded rational curve in \mathbb{CP}^2 is either a line or a smooth conic. Since any line in \mathbb{CP}^2 must have a real point, it follows that C must be a smooth conic.

Conversely, suppose that $C \subset \mathbb{CP}^2$ is a smooth conic without real points. I am going to show that C is the image of a smooth section of τ . To do this, it suffices to show that τ restricts to C to be a diffeomorphism onto \mathbb{S}^* .

Let $[E] \in \mathbb{S}^*$ be an oriented line spanned by the oriented 2-plane $E \subset V$. Consider the real line $\mathbb{P}(E \otimes \mathbb{C}) \subset \mathbb{CP}^2$. Since C is smooth, it cannot have this line as a component and so, by Bezout's theorem [GH, Chapter 1], it must either intersect C transversely in two distinct points or else in a single point of tangency. However, it turns out that a conic without real points cannot have any real tangent lines, so C must intersect each of these real lines $\mathbb{P}(E \otimes \mathbb{C}) \subset \mathbb{CP}^2$ transversely in two points.

Now, the fibers of τ are precisely the connected components of the real lines minus their real points. It follows that τ restricts to C to be a submersion from C to \mathbb{S}^* that is at most 2-to-1. Since both C and \mathbb{S}^* are compact, τ must restrict to C to be a covering space. Finally, since both C and \mathbb{S}^* are diffeomorphic to the 2-sphere, τ must restrict to C to be a diffeomorphism onto \mathbb{S}^* . \square

Remark. The global characterization of smooth rational curves in \mathbb{CP}^2 as plane conics should be thought of as corresponding to the vanishing theorem of §3.6.

4.2. Explicit formulae. In this section, the space of smooth conics in \mathbb{CP}^2 without real points will be used to study the geometry of projectively flat Finsler structures on \mathbb{S} with $K = 1$. Moreover, an explicit formula will be given for the Finsler norm on \mathbb{S} induced by such a conic.

4.2.1. Conics without real points. The natural action of $\mathrm{SL}(3, \mathbb{R})$ on V extends to the complexification of V and to its projectivization \mathbb{CP}^2 as well. This action preserves the projectivization of V itself, namely $\mathbb{RP}^2 \subset \mathbb{CP}^2$. Thus, $\mathrm{SL}(3, \mathbb{R})$ acts on the set of conics without real points. Say that two such conics are \mathbb{R} -equivalent if they are equivalent under this action of $\mathrm{SL}(3, \mathbb{R})$.

Proposition 4. *Fix a basis x, y, z of V^* . Let p and q be real numbers satisfying $|q| \leq p < \pi/2$. Then the conic $C_{p,q} \subset \mathbb{CP}^2$ defined by the equation*

$$(1) \quad e^{ip} x^2 + e^{iq} y^2 + e^{-ip} z^2 = 0$$

is smooth and without real points. Any conic $C \subset \mathbb{CP}^2$ without real points is \mathbb{R} -equivalent to $C_{p,q}$ for some unique $(p, q) \in \mathbb{R}^2$ satisfying $|q| \leq p < \pi/2$.

Proof. Let x, y, z , and (p, q) be as in the statement of the proposition. Write

$$Q = e^{ip} x^2 + e^{iq} y^2 + e^{-ip} z^2 = Q_1 + i Q_2$$

where Q_1 and Q_2 are real quadratic forms on V . Then Q_1 is positive definite because of the assumptions on p and q , so Q cannot vanish on any non-zero element of V . Thus, the conic $C_{p,q}$ that it defines via equation (1) has no real points. Since Q is non-degenerate, $C_{p,q}$ is smooth.

For the remainder of the proposition, let $C \subset \mathbb{CP}^2$ be a conic without real points, defined as the null directions of a complex-valued quadratic form $Q = Q_1 + i Q_2$ where Q_1 and Q_2 are real quadratic forms on V . Note that Q is uniquely determined by C up to a complex scalar multiple.

A slightly messy but straightforward argument shows that unless Q_1 and Q_2 are simultaneously diagonalizable they will have a common (real) null vector, which is impossible since C has no real points. Thus, let x, y , and z be basis elements of V^* so that

$$Q = a_1 x^2 + a_2 y^2 + a_3 z^2$$

where the a_i are complex numbers. The positive real linear combinations of the a_i cannot contain $0 \in \mathbb{C}$ since, otherwise Q would have a non-zero real null vector. By multiplying Q by a non-zero complex number, it can be arranged that the positive linear combinations of the a_i consist of the positive real linear combinations of $\{e^{ip}, e^{-ip}\}$ for some unique real number p satisfying $0 \leq p < \pi/2$. By scaling and rearranging the elements of the basis x, y, z , it can be arranged that $a_1 = e^{ip}$ and $a_3 = e^{-ip}$ while $a_2 = e^{iq}$ for some unique q satisfying $-p \leq q \leq p$. \square

Remarks. When $|q| < p$, the stabilizer of $C_{p,q}$ in $\mathrm{SL}(3, \mathbb{R})$ is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$ and consists of the unimodular transformations of the form $(x, y, z) \mapsto (\pm x, \pm y, \pm z)$. When $0 < |q| = p$, the stabilizer is $\mathrm{O}(2)$. When $p = q = 0$, the stabilizer is $\mathrm{SO}(3)$.

The complex-valued quadratic form

$$Q = e^{ip} x^2 + e^{iq} y^2 + e^{-ip} z^2$$

constructed in the course of the proof of Proposition 4 is uniquely determined by C up to a positive real multiple. A quadratic Q of this form will be said to be *normalized*. One could further require that $x \wedge y \wedge z$ be a unit volume form on V , which would make Q unique, but that will not be needed below.

4.2.2. Finsler structures on \mathbb{S} . In most of this article, a Finsler structure has been defined to be a hypersurface in the tangent bundle of a surface. However, a Finsler structure Σ can also be specified by its corresponding ‘norm’ on the tangent bundle, i.e., the function $|\cdot|_\Sigma$ on the tangent bundle that satisfies $|\lambda v|_\Sigma = \lambda |v|_\Sigma$ for all tangent vectors v and $\lambda \geq 0$ and $|v|_\Sigma = 1$ for all $v \in \Sigma$. (The quotes around ‘norm’ are intended to remind the reader that $|-v|$ is not generally the same as $|v|$.)

In this final section, it will be convenient to describe the Finsler structures of Theorem 9 in terms of their corresponding norms.

As discussed in §4.1, for any pair of vectors $\mathbf{v} = (v, w)$ in $V = \mathbb{R}^3$ with $v \neq 0$, a curve $\gamma_{\mathbf{v}}$ can be defined in \mathbb{S} by the formula

$$(2) \quad \gamma_{\mathbf{v}}(t) = [v + t w].$$

The velocity at $t = 0$ of this curve will be denoted $[v, w]$ and, is, of course, an element of the vector space $T_{[v]}\mathbb{S}$.

The identity $c[v, w] = [v, c w]$ holds for all real numbers c . Moreover,

$$(3) \quad [v, w] = [a v, a w + b v]$$

for all real numbers $a > 0$ and b . Conversely, if $[v', w'] = [v, w]$, then $[v', w'] = [a v, a w + b v]$ for some real numbers $a > 0$ and b .

It follows that a Finsler norm on \mathbb{S} determines and is determined by a function F on $(V \setminus \{\mathbf{0}\}) \times V$ that satisfies

$$(4) \quad F(v, w) = F(a v, a w + b v) \quad \text{for all real } a > 0 \text{ and } b$$

as well as the homogeneity condition

$$(5) \quad F(v, c w) = c F(v, w) \quad \text{for all } c \geq 0.$$

Theorem 10. *Let $C \subset \mathbb{CP}^2$ be a conic without real points and let Q be a normalized quadratic form on $V \otimes \mathbb{C}$ so that $C = \{[v] \mid Q(v) = 0\}$. Let the inner product of two vectors v and w with respect to Q be denoted $v \cdot w$. Set*

$$(6) \quad F_C(v, w) = \Re \left[\frac{\sqrt{(w \cdot w)(v \cdot v) - (w \cdot v)^2} - i(v \cdot w)}{(v \cdot v)} \right].$$

Then F_C defines the Finsler norm of the Finsler structure on \mathbb{S} with linear geodesics and $K = 1$ whose canonical section $\sigma : \mathbb{S}^ \rightarrow \mathbb{CP}^2$ has its image equal to C .*

Remark. Before beginning the proof (which will be a calculation), it probably is a good idea to explain that the quantity inside the radical in the formula for F_C can never be a negative real number and the branch of the complex square root being used is the one satisfying $\sqrt{1} = 1$ and having the negative real axis as its branch locus. Formula (6) then defines a function on $(V \setminus \{\mathbf{0}\}) \times V$ satisfying (4) and (5). It will be seen below that this function is positive and smooth away from points of the form (v, cv) . (These points represent the zero section of $T\mathbb{S}$.)

By the remarks after Proposition 4, the normalized Q is only determined by C up to a positive real multiple. However, replacing Q by rQ for any positive r does not affect the resulting function F_C .

Proof. Let $V_{(2)} \subset V \times V$ be the set of pairs of linearly independent vectors in V . This is a connected, open subset of $V \times V$.

Let (v, w) be an element of $V_{(2)}$. Since C has no real points and no real tangents, the complexified line spanned by v and w must intersect C transversely in two non-real points.

Let $p = [\alpha v + \beta w]$ be such a point. Since it is not real, neither α nor β can vanish, nor can the ratio α/β be real. By multiplying α and β by an appropriate scalar, it can be assumed that $\beta = ib$ for some non-zero real number b . Then α cannot be pure imaginary, so by dividing both α and β by the real part of α , it can be assumed that $\alpha = 1+ia$ for some real number a . Thus, such a $p \in C$ can be uniquely written in the form

$$p = [(1+ia)v + ibw]$$

for some real numbers a and $b \neq 0$.

Now, there are two such points,

$$\begin{aligned} p_1 &= [(1+ia_1)v + ib_1w] = [v + i(b_1w + a_1v)], \\ p_2 &= [(1+ia_2)v + ib_2w] = [v + i(b_2w + a_2v)]. \end{aligned}$$

Since $\tau(p_i) = [v \wedge (b_i w + a_i v)] = [b_i v \wedge w]$ and these two points must represent the two different oriented lines spanned by v and w , it follows that b_1 and b_2 are always of different signs. Thus, it can be arranged that $b_1 < 0 < b_2$. Once this is done, b_1 and b_2 can be regarded as smooth (in fact, analytic) functions on $V_{(2)}$. Note that, while $b_2(v, -w) = -b_1(v, w)$, it will not generally be true that $b_1(v, w) = -b_2(v, w)$ unless the conic C is invariant under conjugation.

By Theorem 9, the Finsler structure Σ_C is defined on \mathbb{S} so that the tangent vector $[v, (b_2 w + a_2 v)] = b_2[v, w]$ is a unit vector. To establish Theorem 10, it suffices to show that the formula for b_2 is just

$$b_2(v, w) = \frac{1}{F_C(v, w)}.$$

This should be just a calculation, but it turns out that some argument is needed to show that the formula F_C makes sense, particularly as regards the choice of sign inherent in the definition of the square root. It is here that having Q be normalized is important.

First, I will show that the quantity under the radical in (6) is never a negative real number and is, in fact, never zero for $(v, w) \in V_{(2)}$. Since C is without real points and Q is normalized, Proposition 4 implies that there exist non-negative real numbers p and q and a basis of V^* , i.e., an identification of V with \mathbb{R}^3 , so that, for any $v = (x, y, z)$, the formula

$$Q(v) = v \cdot v = e^{ip} x^2 + e^{iq} y^2 + e^{-ip} z^2$$

holds. If $w = (a, b, c) \in V$ is any other vector, then, of course,

$$w \cdot w = e^{ip} a^2 + e^{iq} b^2 + e^{-ip} c^2$$

and

$$w \cdot v = e^{ip} ax + e^{iq} by + e^{-ip} cz.$$

Calculation now yields that $(v \cdot v)(w \cdot w) - (w \cdot v)^2$ is equal to

$$e^{i(q-p)} (yc - zb)^2 + (xc - za)^2 + e^{i(q+p)} (xb - ya)^2.$$

With a slight abuse of notation, denote this quantity by $Q(v \wedge w)$. The above expression for $Q(v \wedge w)$ shows that it lies in the wedge $W_{p,q}$ consisting of the non-negative real linear combinations of $e^{i(q+p)}$, 1, and $e^{i(q-p)}$. Since, by hypothesis, $|q| < p \leq \pi/2$, this wedge does not contain any negative real numbers. Thus, $Q(v \wedge w)$ cannot be a negative real number and cannot equal zero unless $(yc - zb) = (xc - za) = (xb - ya) = 0$, i.e., unless v and w are linearly dependent. In particular, taking the complex square root function to be branched along the negative real axis and to satisfy $\sqrt{1} = 1$, the expression $\sqrt{Q(v \wedge w)}$ defines a smooth (in fact, analytic) function on $V_{(2)}$ with values in the right half plane of \mathbb{C} .

Now, for given $(v, w) \in V_{(2)}$, consider the defining equation for p_2 . This is

$$0 = ((1+ia_2)v + ib_2w) \cdot ((1+ia_2)v + ib_2w),$$

Together with the inequality $b_2 > 0$. Expanding this equation, dividing by $v \cdot v \neq 0$ and collecting yields

$$0 = \left(1 + ia_2 + ib_2 \frac{(v \cdot w)}{(v \cdot v)}\right)^2 - b_2^2 \frac{Q(v \wedge w)}{(v \cdot v)^2}.$$

After multiplying by -1 , the right hand side of this equation factors as

$$\left(a_2 + b_2 \left(\frac{(v \cdot w)}{(v \cdot v)} + i \frac{\sqrt{Q(v \wedge w)}}{(v \cdot v)}\right) - i\right) \left(a_2 + b_2 \left(\frac{(v \cdot w)}{(v \cdot v)} - i \frac{\sqrt{Q(v \wedge w)}}{(v \cdot v)}\right) - i\right).$$

Each of these factors is a real analytic function on $V_{(2)}$, so one of them must vanish identically. Accordingly, write

$$(7) \quad a_2 + b_2 \left(\frac{(v \cdot w)}{(v \cdot v)} + i\varepsilon \frac{\sqrt{Q(v \wedge w)}}{(v \cdot v)}\right) = i,$$

where $\varepsilon = \pm 1$. Since a_2 is real and b_2 is real and positive, the quantity in the parentheses must have positive imaginary part. Note that the real and imaginary parts of (7) constitute a pair of linearly independent equations for a_2 and b_2 .

The same argument applied to p_1 yields

$$(8) \quad a_1 + b_1 \left(\frac{(v \cdot w)}{(v \cdot v)} - i\varepsilon \frac{\sqrt{Q(v \wedge w)}}{(v \cdot v)}\right) = i.$$

(The opposite sign must hold since $b_1 \neq b_2$.) Since a_1 is real and b_1 is real and negative, the quantity in the parentheses in (8) must have negative imaginary part.

It now follows from (7) and (8) and the signs of the b_i that the real part of

$$\varepsilon \frac{\sqrt{Q(v \wedge w)}}{(v \cdot v)}$$

must be positive on $V_{(2)}$. Evaluating this expression at $(v, w) = ((1, 0, 0), (0, 1, 0))$ then yields $\varepsilon = +1$.

Thus,

$$a_2 + b_2 \left(\frac{(v \cdot w)}{(v \cdot v)} + i \frac{\sqrt{Q(v \wedge w)}}{(v \cdot v)} \right) = a_1 + b_1 \left(\frac{(v \cdot w)}{(v \cdot v)} - i \frac{\sqrt{Q(v \wedge w)}}{(v \cdot v)} \right) = i.$$

Finally, taking imaginary parts yields

$$b_2 \Re \left[\frac{\sqrt{Q(v \wedge w)}}{(v \cdot v)} - i \frac{(v \cdot w)}{(v \cdot v)} \right] = 1$$

as was to be shown. \square

Remarks. In conclusion, here are a few remarks about the implications of Theorem 10.

First, for F_C defined as in (6),

$$F_C(v, w) - F_C(v, -w) = 2\Im \left[\frac{(v \cdot w)}{(v \cdot v)} \right].$$

It follows that, if the Finsler structure is to be symmetric, then $(v \cdot w)/(v \cdot v)$ must be real for all non-zero $v \in V$ and all $w \in V$. Referring to the formula for a normalized Q , this can only be true if $p = q = 0$, i.e., if Q is real-valued. In this case, formula (6) reduces to the standard formula for the norm of a Riemannian metric and, of course, the Finsler structure is Riemannian. Thus, the only symmetric Finsler structure among these Finsler structures is the Riemannian one.

Second, the proof shows that $\sqrt{Q(v, w)}/(v \cdot v)$ always has positive real part when $v \wedge w \neq 0$. In particular, this implies that $Q(v \wedge w)/(v \cdot v)^2$ is never a negative real number and equals zero only if $v \wedge w = 0$. This fact does not appear to be easy to establish directly, but, using it, the formula for F_C can be rewritten in the form

$$(9) \quad F_C(v, w) = \Re \left[\sqrt{\frac{(w \cdot w)(v \cdot v) - (w \cdot v)^2}{(v \cdot v)^2}} - i \frac{(v \cdot w)}{(v \cdot v)} \right],$$

a form in which the normalization of Q is irrelevant.

Third, on any particular tangent space $T_{[v]}\mathbb{S}$, the Finsler norm has the form

$$\begin{aligned} F_{[v]} &= \Re \left(\sqrt{P + iQ} \right) - L \\ &= \sqrt{\frac{\sqrt{P^2 + Q^2} + P}{2}} - L \end{aligned}$$

where P and Q are real quadratic forms on $T_{[v]}\mathbb{S}$ while L is a real linear form on this space. It follows without difficulty that the unit vectors of the Finsler structure in each tangent space form an algebraic curve of degree 4, or, in certain degenerate cases, of degree 2. In particular, these spaces are not Randers spaces, nor, indeed, any of the other special types of Finsler surfaces considered in [Ma2].

Fourth, for a fixed basis x, y, z of V^* , as (p, q) ranges through the triangle $|q| \leq p < \pi/2$, the corresponding curves $C_{p,q}$ give rise to Finsler norms $F_{p,q}$ on \mathbb{S} that

are inequivalent. This two parameter family must correspond to the two parameter family of inequivalent compact integral manifolds of the Pfaffian system Θ constructed in §3.8.

It seems that, by some principle of analytic continuation, the non-compact integral manifolds must correspond to the partial Finsler structures on \mathbb{S} induced by conics $C \subset \mathbb{CP}^2$ that have one or more real points. It would be interesting to see what the formula F_C gives for such conics.

Fifth, the values (p, q) where $q = \pm p \neq 0$ correspond to curves $C_{p,q}$ that have a one-parameter symmetry group in $\text{SL}(3, \mathbb{R})$. The corresponding Finsler structures must also be invariant under this subgroup and hence are Finsler surfaces of rotation. Thus, there exists a one parameter family of rotationally invariant, projectively flat Finsler structures with $K = 1$.

Even when $|q| < p$, there is still a $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry group of the Finsler structure $F_{p,q}$ (with an index 2 subgroup consisting of the orientation preserving symmetries). These correspond to reflectional symmetries in three ‘axis’ geodesics.

Finally, since $F_C(v, w) = F_C(-v, -w)$ it follows that each of these Finsler structures is invariant under the antipodal map on \mathbb{S} and hence descends to a well-defined Finsler structure on \mathbb{RP}^2 with linear geodesics and Finsler-Gauss curvature equal to 1.

REFERENCES

- [Ak] H. Akbar-Zadeh, *Sur les espaces de Finsler a courbures sectionnelles constantes*, Acad. Roy. Belg. Bull. Cl. Sci. (5) **74** (1988), 281–322.
- [BaChSh] D. Bao, S.-S. Chern, and Z. Shen, *An Introduction to Finsler Geometry*, Springer-Verlag, New York (manuscript in preparation).
- [Be] L. Berwald, *On Finsler and Cartan geometries. III. Two-dimensional Finsler spaces with rectilinear extremals.*, Ann. of Math. (2) **42** (1941), 84–112.
- [Br] R. Bryant, *Finsler structures on the 2-sphere satisfying $K = 1$* , Finsler Geometry, Contemporary Mathematics **196**, Amer. Math. Soc., Providence, RI, 1996, pp. 27–42.
- [Bu] H. Busemann, *Problem IV: Desarguesian spaces*, Mathematical Developments arising from Hilbert Problems, Proc. Symp. Pure Math., vol. 28, Amer. Math. Soc., Providence, RI, 1976, pp. 131–141.
- [Cara] C. Carathéodory, *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*, B. G. Teubner, Leipzig, 1935 (especially §§274–287).
- [Ca1] É. Cartan, *Sur les variétés à connexion projective*, Bull. Soc. Math. France **52** (1924), 205–241 (Reprinted in *Oeuvres Complètes*, partie III, vol. 1, Éditions du CNRS, 1984).
- [Ca2] ———, *Sur un problème d'équivalence et la théorie des espaces métriques généralisés*, Mathematica **4** (1930), 114–136 (Reprinted in *Oeuvres Complètes*, partie III, vol. 2, Éditions du CNRS, 1984).
- [Ch1] S.S. Chern, *On the Euclidean connection in a Finsler space*, Proc. Natl. Acad. Sci. USA **29** (1943), 33–37 (Reprinted in *Shiing-shen Chern: Selected Papers*, vol. II, Springer-Verlag, New York, 1989, pp. 107–111).
- [Ch2] ———, *Local equivalence and Euclidean connections in Finsler spaces*, Science Reports Tsing Hua Univ. **5** (1948), 95–121 (Reprinted in *Shiing-shen Chern: Selected Papers*, vol. II, Springer-Verlag, New York, 1989, pp. 194–212).
- [Da] G. Darboux, *Leçons sur la Théorie générale des Surfaces*, t. 3, Editions Chelsea, Third edition, 1972, Gauthier-Villars, Paris, 1894–1915.
- [Fu1] P. Funk, *Über Geometrien, bei denen die Geraden die Kürzesten sind*, Math. Annalen **101** (1929), 226–237.
- [Fu2] ———, *Über zweidimensionale Finslersche Räume, insbesondere über solche mit geradlinigen Extremalen und positiver konstanter Krümmung*, Math. Zeitschr. **40** (1936), 86–93.
- [Fu3] ———, *Eine Kennzeichnung der zweidimensionalen elliptischen Geometrie*, Österreichische Akad. der Wiss. Math. Sitzungsberichte Abteilung II **172** (1963), 251–269.

- [GaWi] R. Gardner and G. Wilkens, *Classical geometries arising in feedback equivalence* (1993), Fields Institute, Waterloo (preprint).
- [GH] P. Griffiths and J. Harris, *Principles of Algebraic Geometry*, John Wiley & Sons, Inc., New York, 1978.
- [Ko] S. Kobayashi, *Theory of connections*, Ann. Math. Pura Appl. **43** (1957), 119–194.
- [KN] S. Kobayashi and T. Nagano, *On projective connections*, J. Math. Mech. **13** (1964), 215–236.
- [Ma1] M. Matsumoto, *Projectively flat Finsler spaces of constant curvature*, J. Natl. Acad. Math. India **1** (1983), 142–164.
- [Ma2] ———, *Foundations of Finsler Geometry and Special Finsler Spaces*, Kaiseisha Press, Shigaken, 1986.
- [Po] A. V. Pogorelov, *Hilbert's Fourth Problem*, Winston & Wiley, New York, 1982.
- [Ru] H. Rund, *The Differential Geometry of Finsler Surfaces*, Grundlehren der Math. Wiss., Band 101, Springer-Verlag, Berlin, 1959.
- [Sz] Z. I. Szabó, *Hilbert's fourth problem, I*, Adv. in Math. **59** (1986), 185–301.
- [Un] A. Underhill, *Invariants of the function $F(x, y, x', y)'$ in the Calculus of Variations*, Trans. Amer. Math. Soc. **9** (1908), 316–338.

DEPARTMENT OF MATHEMATICS, DUKE UNIVERSITY, PO Box 90320, DURHAM, NC 27708-0320

E-mail address: bryant@math.duke.edu